PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/00, 19/10, 19/06	A1	(11) International Publication Number: WO 98/03532
		(43) International Publication Date: 29 January 1998 (29.01.98)

(21) International Application Number: PCT/US97/12888

(22) International Filing Date: 23 July 1997 (23.07.97)

(30) Priority Data:
60/022,573   24 July 1996 (24.07.96)   US

(71)(72) Applicant and Inventor: HANNA, Michelle, M. [US/US]; 804 Hoover Street, Norman, OK 73072 (US).

(74) Agents: PALMER, John et al.; Ladas & Parry, Suite 2100, 5670 Wilshire Boulevard, Los Angeles, CA 90036 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: BASE-PROTECTED NUCLEOTIDE ANALOGS WITH PROTECTED THIOL GROUPS

(57) Abstract

The present invention is directed to protected thiol analogs of pyrimidine bases for syntheses of DNA and RNA by chemical or enzymatic methods. The subject analogs include reagents suitable for DNA or RNA synthesis via phosphoramidite, H-phosphonate or phosphotriester chemistry as well as reagents suitable for use by RNA and DNA polymerase, including thermostable polymerases employed by PCR or other nucleic acid amplification techniques. The nucleotide analogs synthesized by methods of this invention can thus be incorporated into oligonucleotides or polynucleotides, deprotected and derivatized with a functional group. In some cases the protecting groups are themselves antigenic and may be left on the oligonucleotides or polynucleotides for detection with antibodies. A method of synthesizing oligonucleotides with a functional group using the subject nucleotide analogs is also provided.

BNSDOCID: <WO___9803532A1_I_>

BASE-PROTECTED NUCLEOTIDE ANALOGS WITH PROTECTED THIOL GROUPS

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

Not Applicable

BACKGROUND OF THE INVENTION

The present application claims the benefit of the filing date of U.S. Provisional Application 60/022,573, filed July 24, 1996.

The present invention is directed to thiol-protected pyrimidine nucleotide analogs which can be used, as one example, for syntheses of DNA and RNA by chemical or enzymatic methods. The subject analogs include reagents suitable for DNA or RNA synthesis via phosphoramidite, H-phosphonate or phosphotriester chemistry as well as reagents suitable for use by RNA and DNA polymerases, including thermostable polymerases employed by PCR or other nucleic acid amplification techniques. Methods of synthesizing the nucleotide analogs are also provided by the present invention. The nucleotide analogs of this invention can thus be incorporated into oligonucleotides or polynucleotides, deprotected, and then derivatized with a functional group.

Oligonucleotides with a variety of modifications have widespread utility for many purposes, such as stabilizing oligonucleotides to degradation, introducing reporter groups, allowing site-specific delivery of therapeutics, and introducing crosslinkers. Such modifications can occur as modified internucleotide phosphate linkages or analogs of such linkages,

1

modified sugars or modified bases. Additionally, 5'- or 3'-end conjugates of the oligonucleotides represent another class of modified oligonucleotides. The present invention relates to base-modified nucleotide analogs with protected thiol groups; these analogs are intermediates for chemical or enzymatic synthesis of oligonucleotides and polynucleotides.

The synthesis of oligoribonucleotides and oligodeoxyribo-nucleotides (i.e. oligonucleotides) containing base-modified nucleotides at specific positions provides a powerful tool in the analysis of protein-nucleic acid or nucleic acid-nucleic acid interactions. These oligonucleotides have many other potential uses, such as the site-directed delivery of therapeutics, utility as anti-sense therapeutics, and utility as diagnostic probes. Nucleotide analogs can be introduced into nucleic acids either enzymatically, utilizing DNA and RNA polymerases, or chemically, utilizing manual or automated synthesis. Preparation of such oligonucleotides by automated synthesis utilizing, for example, phosphoramidite nucleotides allows for incorporation of a broad range of nucleotide analogs without the restraints for specific substrate conformation of the nucleotides that is imposed by most polymerases. Often, nucleotide analogs containing photoreactive crosslinking groups are introduced into oligonucleotides to probe protein-nucleic acid interactions via photocrosslinking (for a partial review, see Hanna, 1989, Methods Enzymol. 180:383-409; and Hanna, 1996, Methods Enzymol, 273 Chapter 31). Deoxyoligonucleo-tides containing 4-thiothymidine have been prepared and used for

-2-

photochemical crosslinking of proteins directly to the nucleotide bases through the group (Nikiforov et al., 1992, Nucleic Acids Res. 20:1209-14). Similarly, oligonucleotides containing 3-(amino-propyl)-2'-deoxyuridine have been prepared and the amino group subsequently modified with fluorescent, photoactive or other reporter groups (Gibson et al., 1987, Nucleic Acids Res. 15:6455-66). However, the former thiodeoxynucleotide suffers the disadvantage that if it is modified with a thiol modifying reagent the normal Watson-Crick base of the nucleotide is drastically affected, making this and similar analogs generally unsuitable for use in enzymatic nucleic acid synthesis.

Other analogs involving modifications at the C5 position of deoxyuridine have also been previously reported. The thiol-containing analog, 5-thiocyanatodeoxyuridine phosphoramidite, provided a 5-mercaptodeoxyuridine moiety within the oligonucleotide following reduction of the thiocyanate (Bradley & Hanna, 1992). However, the thiocyanato moiety displayed variable stability during synthesis of both the nucleotide and the oligonucleotide and therefore this compound did not represent an ideal analog for incorporation of 5-thiol modified nucleotides into nucleic acids. The syntheses of a series of phosphoramidites containing alkylthiol tethers at the C5 position of deoxyuridine has been reported (Goodwin & Glick, 1993). The thiol groups in these analogs are attached to the ring by either a three, four, or five carbon chain. The presence of the carbon chains makes the minimal distance between the molecular probes and the oligonucleotide greater than

-3-

that which can be achieved with our analog. In addition, these
compounds represent alkyl thiol analogs which have a lower
reactivity for modification of the thiol group than the 5-
mercaptopyrimidine analog. This is due to an increase in acidity

5   of 5-mercaptopyrimidines (pKa~5-5.6) over alkylthiol moieties
(pKa-8-10). Phosphoramidites containing alkylthiol tethers at the
N3 position of thymidine have also been prepared, but the position
of this modification results in a disruption of the Watson-Crick
base pairing. These analogs have been used mainly for preparing

10   disulfide cross-links in DNA for studies involving stem loop and
triple helical structures (Glick, 1991; Goodwin et al., 1994).

Several 5-modified deoxyuridine phosphoramidites are
commercially available which contain functional groups (i.e.
carboxylic acids or alkyl amines) for post-modification following

15   incorporation into the oligonucleotide. Likewise, alkyl thiol
ethers of 5-mercaptodeoxyuridine containing protected carboxylic
acids and alkyl amines have also been described (Bergstrom et al.,
1991). The protected functional groups in these analogs are not
attached directly to the ring but are positioned at the end of

20   carbon chains. These groups are not as easily modified as
mercaptans: the functional groups formed during post-synthetic
modification are limited, and an easily cleavable group is not
available. Thiol-containing phosphoramidites for incorporating 4-
thiothymidine (Clivio et al., 1992; Xu et al., 1992c), 4-

25   thiodeoxyuridine (Clivio et al., 1992; Coleman & Kesicki, 1994), 2-
thiothymidine (Connolly & Newman, 1989; Kuimelis & Nambiar, 1994),

-4-

6-thiodeoxyguanosine (Christopherson & Broom, 1991; Waters & Connolly, 1992; Xu et al., 1992b) and 6-thioinosine (Clivio et al., 1992a) have been reported. These analogs can occupy internal positions within an oligonucleotide and can serve as photochemical crosslinkers. However, they cannot be further modified without disrupting Watson-Crick base pairing, and therefore, as photocrosslinking probes, these analogs are only useful for evaluating interactions which occur directly with the nucleotide base. Modification with other molecular probes (e.g., fluorescent tags) would also disrupt Watson-Crick base pairing. In addition, the deprotection of oligonucleotides containing these thiol-modified nucleotides must be carefully monitored to prevent conversion of these analogs to the corresponding oxygen and nitrogen derivatives.

Described herein are nucleotide analogs which can be used for site-specific modification of DNA or RNA at internal and terminal positions within the DNA or RNA sequence, and after modification, for molecular probes which can be placed at variable distances from the DNA or RNA backbone.

The present invention provides novel base-protected nucleotide analogs, both ribonucleotides and deoxynucleotides, that contain masked thiol groups on the 5 position of pyrimidines, which is not involved in Watson-Crick base pairing. These analogs can be incorporated into oligonucleotides via automated synthesis and isolated with the thiol protecting group intact. After removal of the thiol protecting group many types of functional groups, such as

-5-

photocrosslinking agents, fluorescent tags, radioisotopes, biotin, reporter molecules, spin labels (e.g., commercially available proxyl or tempo), chemiluminescent, antigenic or other functional groups, can be site-specifically attached by utilizing thiol-

5   modifying reagents. This feature adds a level of specificity to the oligonucleotide modifications not present with the amino-tagged analogs previously described (Gibson et al, 1987), and enables examination of molecular interactions that are not directly at the nucleotide base by allowing functional groups to be placed at

10   varying distances from the base or helix strand. Since these analogs have the functional group attached via the sulfur atom, some have the further advantage of being cleavable under conditions which will not degrade or modify the oligonucleotide.

SUMMARY OF THE INVENTION

15   This invention relates to pyrimidine nucleotide analogs which contain modified bases with protected thiol groups attached at a position on the base, preferably the 5 position, which is not involved in Watson-Crick base pairing. These nucleotide analogs are intermediates in chemical or enzymatic synthesis of DNA or RNA

20   oligonucleotides and are therefore stable under conditions required for synthesis of these molecules. After synthesis, the protecting group on the analog is removable to generate a reactive thiol group. Once generated, the thiol group can be treated with thiol modifying reagents to attach functional groups such as crosslinking

25   agents or reporter molecules.

-6-

In particular, the nucleotide analogs of the present invention have the formula:

wherein

R$_1$ is −H, −OH, a mono, di, or triphosphate group, or
    −OR$_4$;

R$_2$ is −H, −OH, a mono, di, or triphosphate group, a
    phosphoramidite group, a phosphorothioamidite
    group, a phosphonate group, an O-substituted
    monophosphate group, −OR$_4$, or a solid support bonded
    via an O at the 3' position;

R$_3$ is −H, −OH, a mono, di, or triphosphate group, or
    −OR$_4$;

R$_4$ is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected
    thiol group attached at the 5 position on said base
    that is not involved in Watson-Crick base pairing
    or does not disrupt normal Watson-Crick base
    pairing, said protected thiol group being stable
    under conditions of chemical nucleic acid synthesis
    and/or conditions of enzymatic nucleic synthesis
    and being convertible to a reactive thiol (SH)
    after said synthesis.

−7−

Preferred nucleotide analogs of the present invention are the protected phosphoramidites or 5' mono, di and triphosphates of modified cytosine or uridine bases for use in the chemical synthesis of DNA and RNA by the phosphoramidite method or enzymatic synthesis with polymerases, ligases, or other nucleotide polymerizing enzymes.

In addition, the nucleotide analogs of the present invention include other nucleoside phosphates, containing 3', 5', or 3',5' monophosphates, diphosphates, or triphosphates and further optionally comprising 2' mono, di, or triphosphates when the nucleoside is a ribonucleoside.

Another aspect of this invention provides nucleic acids and oligonucleotides containing the subject nucleotide analogs having a protected thiol group on a base moiety of that nucleic acid or oligonucleotide. A method is also provided to synthesize these nucleic acids or oligonucleotides, deprotect the thiol group and attach a functional group to the reactive thiol moiety.

Yet another aspect of this invention is directed to a method of synthesizing the subject base-protected nucleotide analogs.

DESCRIPTION OF THE DRAWINGS

Figure 1: Synthetic pathway for the DNP-labeled phosphoramidite (compound 7). Reaction conditions are as follows a) $SCNCl, CH_3CO_2H$; b) $DTT, MeOH, EDTA$; c) 2,4-dinitrofluorobenzene, Et_3N, CH_3CN; d) $NaOMe, MeOH$; e) $DMT^+BF_4^-, DBMP, CH_3CN$; f) 2-

-8-

cyanoethyl-N,N,N',N'-tetraisopropyl-phosphorodiamidite,tetrazole,CH₃CN.

Figure 2:   A. RP HPLC chromatograms of Oligonucleotides A and B.   The purified 7-mers were analyzed by analytic HPLC on an Ultrasphere ODS column, eluted with a 15 minute linear gradient from 5% to 20% (v/v) acetonitrile in 20 mM TEAA (pH 7.1). (a) Oligonucleotide A, GTA TGT A, eluted at 9.8 min.   (b) DNP-labeled oligonucleotide B, GTA T*GT A, eluted at 11.6 min, where T represents incorporation of the DNP-analog described in Figure 1. B.RP HPLC profile of enzymatic degradation of Oligonucleotides A and B.   The 7-mers were subjected to enzymatic treatment with Nuclease P1 and CIP.   The reaction mixture was analyzed by analytical HPLC on an Ultrasphere ODS column, eluted with a linear gradient of acetonitrile in 100 mM TEAA (pH 7.1) from 0% to 5% (v/v) in 5 min, 5% to 6% (v/v) in 15 min., 6% to 50% (v/v) in 5 min and finally 50% to 80% (v/v) in 15 min.   (a) oligonucleotide A reaction.   (b) DNP-labeled Oligonucleotide B.

Figure 3:   A.: Gel electrophoretic analysis of DNP-labeled oligonucleotides.   The[³²P] labeled oligonucleotides C-F (Table 1) were analyzed by electrophoresis on a 25% polyacrylamide-7M urea gel (acrylamide/methylene bisacrylamide=19:1).   Lane 1: Commercial 24-mer marker, CGC GGA ATT CTC ATG CAT TGC CCA;   Lane 2; 24-mer, Oligonucleotide E;   Lane 3: DNP-labeled 24-mer, Oligonucleotide F; Lane 4: 22-mer, Oligonucleotide C;   Lane 5: DNP-labeled 22-mer, Oligonucleotide D.   B.: Gel electrophoretic analysis of the deprotection and alkylation reactions.   Oligonucleotides E and F

-9-

were reacted with 1.4 M BME. Half of these reactions were alkylated with 5-IAF. The treated oligonucleotides were analyzed by electrophoresis on a 25% polyacrylamide-7M urea gel (acrylamide/methylene bisacrylamide=19.1). Lane 1: Oligonucleotide E + 1.4 M BME at room temperature; Lane 2: Lane 1 reaction + 5-IAF; Lane 3: Oligonucleotide E + 1.4 M BME at 45°C; Lane 4: Lane 3 reaction + 5-IAF; Lane 5: Oligonucleotide F + 1.4 M BME at room temperature; Lane 6: Lane 5 reaction + 5-IAF; Lane 7: Oligonucleotide F + 1.4 M BME at 45°C; Lane 8: Lane 7 reaction + 5-IAF.

Figure 4: Analysis of PCR products.

A. PCR products were analyzed by electrophoresis on a 1.5% agarose gel. DNA was detected with ethidium bromide. Lane 1: λ Hind III markers; Lane 2: PCR product using Oligonucleotide E as a primer; Lane 3: PCR product using DNP-labeled Oligonucleotide F as a primer. XC represents xylene cyanol and BPB represents bromophenol blue.

B. Dot Blot immunodetection of DNP-labeled DNA. PCR products were attached to a nylon membrane, incubated with a primary antibody to DNP and a 2° antibody linked to horseradish peroxidase. Chemiluminescence was used for detection. Upper panel: PCR product using DNP-labeled Oligonucleotide F as a primer; Lower panel: PCR product using Oligonucleotide E as a primer.

Figure 5: Protein Crosslinking to Site-Specifically Modified DNA

To analyze the DNA-protein complexes in the transcription complex a deoxynucleoside phosphoramidite was incorporated site-

-10-

specifically into a single-stranded DNA oligonucleotide via automated synthesis. In these studies, the nucleotide analog was modified with a masked reactive thiol group on the non-basepairing 5 position of deoxyuridine (5-S-DNP-dU, Figure 1). The oligonucleotide was then radioactively labeled uniquely on the 5' phosphate of the nucleotide analog with polynucleotide kinase and $[\gamma\text{-}^{32}P]$ ATP (Step 1). The radioactively labeled, base-modified oligonucleotide was then annealed to a single-stranded DNA template (Step 2) and incorporated site-specifically into double stranded DNA by primer extension (Step 3). The protecting group was then removed from the double-stranded DNA by treatment with β-mercaptoethanol (Step 4), and after removal of the BME, a functional group was added to the unmasked thiol group. For crosslinking, the DNA was reacted with azidophenacyl bromide to attach a photoreactive azide group (Step 4). RNA polymerase and transcription factors are then added to assemble a complex on the modified DNA, and the complex is irradiated with ultraviolet light to covalently join the DNA to proteins which are bound adjacent to the crosslinker (Step 5). Lastly, the DNA-protein complex is treated with a nuclease which leaves 5' NMPs to digest the DNA to a small fragment (Step 6). This will leave a single radioactive label on the 5' phosphate of the nucleotide involved in the DNA-protein crosslink.
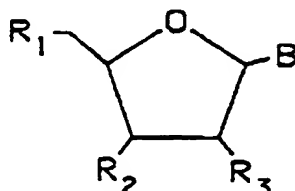
-11-

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a series of novel nucleotide analogs which are masked synthons for use as intermediates in chemical or enzymatic synthesis of nucleic acids, including synthesis of both oligonucleotides and polynucleotides. The nucleotide analogs of this invention, which contain a protected thiol group, can thus be incorporated into DNA or RNA under standard synthetic conditions without loss of the thiol protecting group. This stability of the thiol protecting group permits site-selective introduction into a nucleic acid of the nucleotide analog in a manner which facilitates later addition of a functional group at that site. Thus, the subject oligonucleotides (or polynucleotides) can contain one (or more) of the subject nucleotide analogs.

Thus, functional groups which can not withstand the conditions for chemical nucleic acid synthesis, especially during automated synthesis, or which may be too bulky or sterically hindered for enzymatic nucleic acid synthesis can be readily incorporated into the final oligonucleotide (or polynucleotide) product. Such derivatized nucleic acids have great utility in studying protein-nucleic acid or nucleic acid-nucleic acid interactions, as well as the potential for diagnostics and delivery of therapeutics. Hence, the present invention permits the skilled artisan to place chemical tags such as crosslinking groups, fluorescent molecules, radioisotopes or other reporter molecules at specific positions on nucleic acid molecules for analysis of molecular mechanisms, for

-12-

creation of diagnostic probes, for therapeutics, for antisense therapeutics and for many other purposes.

In particular, the nucleotide analogs of the present invention are intermediates for the chemical synthesis of DNA and RNA by manual or automated techniques and are represented by the formula:

wherein

R_1 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R_2 is -H, -OH, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, -OR$_4$, or a solid support bonded via an O at the 3' position;

R_3 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R_4 is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at a position on said base, preferably the 5 position, that is not involved in Watson-Crick base pairing or does not disrupt normal Watson-Crick base pairing, said protected thiol group being stable under conditions of

-13-

chemical nucleic acid synthesis and/or conditions of enzymatic nucleic synthesis and being convertible to a reactive thiol after said synthesis.

5   As used herein, protecting groups defined by R_4 are known to those of ordinary skill in the art and include any known protecting group suitable for protection of the 2', 3' or 5' hydroxyls of ribose and the 3' or 5' hydroxyls of deoxyribose sugars. In this regard, Greene et al. (1990) <u>Protecting Groups in Organic</u>
10   <u>Synthesis</u>, <u>2nd ed.</u>, John Wiley & Sons, Inc., New York, NY, provides a comprehensive review of protecting groups, and methods of preparing the corresponding protected compounds, which can be used for different reactive groups, including reactive hydroxyl groups.

         Accordingly, examples of protecting groups defined by R_1, R_2
15   and R_4 are lower alkyl, lower cyanoalkyl, lower alkanoyl, aroyl, aryloxy, aryloxy-lower alkanoyl, haloaryl, fluorenylmethyloxy-carbonyl (FMOC), levuloyl, 9-phenylxanthene-9-yl, trityl, p-monomethoxytrityl (MMTr), p-dimethoxytrityl (DMTr), isopropyl, isobutyl, 2-cyanoethyl, acetyl, benzoyl, phenoxyacetyl, halophenyl,
20   1-(2-chloro-4-methylphenyl)-4-methoxy-4-piperidinyl, 2'-acetal, <u>O</u>-nitrobenzyl, <u>tert</u>-butyldimethylsilyl (TBDMS), tetrahydrofuranyl, 4-methoxytetrahydropyranyl and related groups.

         Preferred R_1 protecting groups include 5'OH protecting groups especially DMTr, MMTr, FMOC, levuloyl and 9-phenylxanthene-9-yl
25   groups, and most preferably DMTr. Preferred R_2 protecting groups

-14-

include 3'OH protecting groups, especially the acetyl group. Preferred R_4 protecting groups include 2'OH protecting groups, especially 2'-acetal, tetrahydrofuranyl, 4-methoxytetrahydropyranyl and 1-(2-chloro-4-methylphenyl)-4-methoxy-4-piperdinyl groups, and most preferably TBDMS.

As defined herein, solid supports include controlled pore glass (CPG), polystyrene silica, cellulose, nylon, and the like. Preferred solid supports are CPG and polystyrene. An especially preferred solid support is CPG.

R_2 can be taken with the oxygen atom to which it is attached, to form a phosphoramidite, phosphorothioamidite, a phosphonate, an o-substituted monophosphate, or any other group compatible with chemical nucleic acid synthesis, especially automated DNA or RNA synthesis. As used herein, the phosphoramidite and phosphorothioamidite groups have the general formulas I and II, respectively:

I   II

wherein R_6 is lower alkyl, n-cyano alkyl, substituted lower alkyl, aryl, aralkyl, substituted aralkyl and the like, and R_7 and R_8 are independently lower alkyl or when taken together with the nitrogen to which they are attached comprise cyclic groups including:

-15-

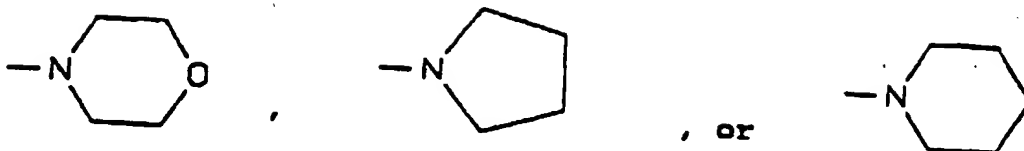preferably the R_6 group of the phosphoramidite group (Formula I) is 2-cyanoethyl (CED) or methyl. Accordingly, the preferred phosphoramidites of this invention are also referred to as CED phosphoramidites or methyl phosphoramidites.

As used herein, the term lower alkyl, when used singly or in combination, refers to alkyl groups containing one to six carbon atoms. Lower alkyl groups can be straight chain or branched and include such groups as methyl, ethyl, propyl, isopropyl, n-butyl, sec-butyl, isobutyl, t-butyl, pentyl, isopentyl, neopentyl, hexyl and the like. The preferred alkyl groups contain one to four carbon atoms.

The term aryl, when used alone or in combination, refers to an aromatic ring containing six to ten ring carbon atoms. The aryl group includes phenyl, and 1- or 2-naphthyl. The preferred aryl group is phenyl.

The term aralkyl refers to aryl groups as described above to which substituents are attached to the aryl by an alkylene bridge. The most preferred aralkyl group is benzyl.

When R_7 and R_8 are lower alkyl, the preferred groups are each isopropyl groups.

-16-

As used herein a phosphonate group (or H-phosphonate) is represented by the formula:

$$O=\overset{\displaystyle \overset{|}{O}}{\underset{\displaystyle \underset{|}{O^-}}{P}}-H$$

and may also be provided as salts, and preferably as triethylammonium salts.

5   As used herein, O-substituted monophosphates have the formula:
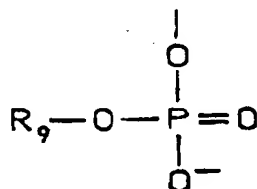
$$R_9-O-\overset{\displaystyle \overset{|}{O}}{\underset{\displaystyle \underset{|}{O^-}}{P}}=O$$

wherein R_9 is lower alkyl, haloalkyl, aryl, haloaryl, or heteroaromatic. By haloalkyl or haloaryl is meant alkyl or aryl groups, respectively, which have been substituted with one or more halogen atoms including F, Cl, Br, or I. Preferred halo

10   substituents are Cl and Br. Preferred R_9 groups include 2-chlorophenyl, 2,5-dichlorophenyl, 2,2,2-trichloroethyl and the N-oxide of 4-methoxypyridine-2-methylene groups.

Modified pyrimidine bases have a protected thiol group attached at a position on the base which is not involved with

15   Watson-Crick base pairing or which does not disrupt normal Watson-Crick base pairing. The protected thiol groups of this invention

-17-

are stable, i.e. not removable, under the conditions used for chemical synthesis of DNA or RNA, and particularly, under the conditions employed in automated DNA or RNA synthesis. Furthermore, the protecting group of the thiol is removable under conditions which do not disrupt the integrity of the oligonucleotide or polynucleotide. In other words, after a nucleotide analog of the present invention has been incorporated into an oligonucleotide, for example, the protected thiol group can be converted to a reactive thiol (SH) to which functional groups can subsequently be added using thiol-modifying reagents.

In accordance with this invention the protected thiol groups include, but are not limited to, the dinitrophenyl group. As used herein, Watson-Crick base pairing refers to the hydrogen bonding pattern of adenine-thymine (AT) base pairs, adenine-uracil (AU) base pairs, or of guanine-cytosine (GC) base pairs. Accordingly, the preferred protected thiol position on pyrimidines is the 5 position on the pyrimidine (C,U) ring.

The preferred bases of the present invention are thus pyrimidines such as cytosine (C), and uracil (U). Cytosine <u>bases</u> contain an additional reactive group, specifically an exocyclic amine, which must also be protected during assembly of the nucleotide chain via chemical synthesis. Accordingly, bases of this invention can have additional protecting groups attached as needed to any of the ring positions. Such protected bases are well known in the art and include but are not limited to, N^2, N^4-anisoyl cytosine, N^4-benzoyl cytosine, and N^4-isobutyryl cytosine. The N^4-

-18-

isobutyryl cytosine is known by the trade name FOD™ base protection (Applied Biosystems, Inc.). Use of these further protected bases is compatible with incorporation of the protected thiol groups as well as subsequent deprotection reactions.

5   The bases of this invention also include any related base analog that is capable of base pairing with a guanine or adenine, the corresponding protected analogs as set forth above for use in chemical synthetic methods to produce DNA and RNA. For example, such base analogs include, but are not limited to, pseudocytosine,

10   isopseudocytosine, 4-acetylcytosine, 2'-O-methylcytosine, dihydrouracil, 2'-O-methyluracil, 2'-O-methyl-pseudouracil, 1-methylpseudouracil, 3-methylcytosine. Bases attached to a ribose or deoxyribose sugar in an α-anomeric configuration can also be present.

15   In a preferred embodiment the bases of the invention are C, U, and N^4 protected C.

The compounds of the present invention also include nucleoside phosphates of the formula:
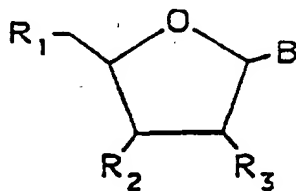
wherein:

20   R_1 is -H, -OH, a mono, di, or triphosphate group, or
                -OR_4;

-19-

R_2 is -H, -OH, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, -OR$_4$, or a solid support bonded via an O at the 3' position;

R_3 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R_4 is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at the 5 position on said base.

In the case of the modified bases the protected thiol group is also stable, i.e. not removable, under the conditions used for enzymatic synthesis of DNA or RNA. Additionally, the pyrimidine bases generally do not require protection of the additional reactive groups such as the exocyclic amines for enzymatic incorporation into DNA or RNA. Accordingly, the preferred bases for this class of compounds includes C and U.

In accordance with this invention the phosphate groups of R_1, R_2 and R_3 embody all the phosphorylated forms at the C_3' and C_5' positions of the sugar moiety (and C_2' position when the sugar is ribose) and include monophosphates, diphosphates, triphosphates and tetraphosphates. Preferably the phosphate group is a triphosphate for R_1.

The preferred compounds of this invention include the phosphoramidites and 5'triphosphates of 5-SDNP-dU, 5-SDNP-U, 5-DNP-

-20-

dC, and 5-SDNP-C. In the case of the phosphoramidites, the base
moieties can contain additional protecting groups on the exocyclic
amines. The preferred compounds also include the 5'triphosphates
of 5-S-S-Et-dU, 5-S-S-Et-U, 5-S-S-Et-dC and 5-S-S-Et-C, 5-S-CH$_2$-

5   NHCOCH$_2$Ph-dU,   5-S-CH$_2$-NHCOCH$_2$Ph-U,   5-S-CH$_2$-NHCOCH$_2$Ph-C,   5-S-CH$_2$-
NHCOCH$_2$Ph-dC, 5-SCH$_2$CH$_2$DNP-dU, 5-SCH$_2$CH$_2$DNP-U, 5-SCH$_2$CH$_2$DNP-dC, and
5-SCH$_2$CH$_2$DNP-C.

The nucleotide analogs of the present invention can be
prepared by adding a protected thiol group to the base moiety of

10   the desired nucleoside.   The so-modified nucleoside can then be
phosphorylated to produce a nucleotide analog phosphate compound of
this invention using conventional phosphorylation techniques.   To
produce a phosphoramidite, phosphonate, phosphorothioamidite or O-
substituted monophosphate of this invention, the 5'-OH or 2'-OH

15   groups of above-modified nucleoside are protected as necessary by
addition of the desired protecting group by standard methodology.
This protection step(s) is (are) followed by conversion to the
nucleotide-analog   phosphoramidite,   phosphorothioamidite,
phosphonate or O-substituted monophosphate by reaction of the 3'OH

20   of the nucleoside with the appropriate modifying group. Similarly,
the 3'OH can be attached to a solid support, such as CPG, or
another protecting group using conventional methodology available
to the ordinarily skilled artisan in this field.

For example, to prepare protected thiol-groups at the 5

25   position of uridine or 2'-deoxyuridine nucleosides,the sulfur atom
must be incorporated on the base, for which the method of Nagamachi

-21-

et al. (1972, J.C.S. Chem. Commun. 18:1025-6) has been modified. An equivalent amount of potassium thiocyanate is dissolved into a solution of chlorine gas in glacial acetic acid to produce the thiocyanogen chloride in solution. An iodometric titration is

5   performed with Cl_2/HOAc to determine the amount of Cl_2 present. The solution can be used directly or filtered prior to the next step. In either case, the acetyl-protected U or dU is added all at once with stirring. This reaction is maintained with stirring at room temperature for about 1.5 hours until the reaction is completed.

10   An excess of cyclohexene or other quenching agent is then added to quench any remaining thiocyanogen chloride. Quenching is complete in 15 to 60 min, and usually in about 30 min. After removing solvents and organic residue in vacuo, the remaining residue can be purified by silica gel chromatography with a gradient of petroleum

15   ether and ethyl acetate or with other chromatographic methods. These acetyl-protected 5-thiocyanato dU and U nucleosides can then be reduced to 5-mercapto nucleosides for further derivatization as described herein to produce the nucleotide analogs of the present invention.

20   The mono-, di-, tri or tetra phosphate compounds can be prepared by a similar reaction by beginning with the appropriate starting 5-thioprotected nucleoside. If necessary, the 2'-OH of the ribose sugar can be protected prior to the reaction using any of the known 2'-OH protecting groups by conventional techniques.

25   The described synthesis and utilization of 5-thiocyanato-deoxyuridine phosphoramidite for placement of the sulfur directly

-22-

on the base has been previously described (Bradley & Hanna, 1992).
However, the synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-S-(2,4-
dinitrophenyl) mercapto-2'-deoxyuridine-3'-O-(2-cyanoethyl-N,N'-
diisopropyl) phosphoramidite is more reproducible and the product

5   is more stable than 5-thiocyanatodeoxyuridine phosphoramidite and
represents a more desirable analog. A variety of functional groups
can be attached to the reactive thiol. Once this modified
nucleotide is incorporated into an oligonucleotide or nucleic acid,
the protecting group can be removed to unmask the reactive thiol,

10   making it available for further derivitization with thiol-specific
agents. This method has the advantage that an oligonucleotide
tagged with this analog can be synthesized and stored for long
periods without removal of the protecting group. When needed, an
aliquot can then be deprotected and modified. In addition, in some

15   cases the protecting groups are themselves antigenic (e.g., ⁻DNP)
and may be retained in the nucleic acid product for detection
utilizing antibodies.

One method used to produce one of the nucleotide analog
phosphoramidites of this invention and described in detail in

20   Example 1 below is shown in Fig. 1. Briefly, as depicted, 5-SDNP-
dU or 5-SDNP-U are prepared by reduction of the SCN, protection of
the thiol as the DNP ether, and de-esterification of the sugar
ring. The protected-thiol nucleoside thus formed is then reacted
with the DMT⁺BF₄⁻, or another 5'OH protecting group, in an anhydrous

25   organic solvent in the presence of an organic base for 2-24 hours
until the reaction is complete. The resulting product can be

-23-

isolated by liquid chromatographic methods and then converted to a CED phosphoramidite by reaction with 2-cyanoethyl N, N, N', N'-tetraisopropyl phosphordiamidite and tetrazole under anhydrous conditions and in an inert atmosphere. This reaction is preferably

5   stirred for 1 hour, although this time can be varied, the solvents removed in vacuo and the residue purified by chromatographic methods. Exposure to atmosphere is acceptable but should be minimized. This phosphoramidite can be stored dry under positive pressure of inert atmosphere (argon or nitrogen) in a tightly

10   sealed container at low temperature.

Other phosphoramidites of this invention can be prepared by reaction with the appropriate chlorophosphoroamidite. For example, the O-methylphosphoramidites can be prepared by reacting the 5-SDNP-5'-DMTr-dU with N,N-diisopropylmethylphosphonamidic chloride

15   by conventional techniques. Similarly, the phosphorothioamidite, phosphonates, O-substituted monophosphate can be prepared using conventional techniques with commercially available reagents. As mentioned above to avoid unwanted side reactions, protection of the 5'OH and 2'OH groups on the sugar moiety and the exocyclic amine

20   groups on the base moiety may be necessary before the final reaction step which produces the nucleotide analogs of this invention.

In general the chemical synthetic routes to nucleotide phosphoramidites, phosphorothioamidites, phosphonates and O-

25   substituted monophosphates, as well as nucleotide phosphates, are well known. In addition, nucleoside phosphates can be

-24-

enzymatically synthesized. Chemical synthetic techniques for these compounds, as well as the common synthetic routes to prepare RNA and DNA, have been described in many sources. Particularly, useful references include Gait (1984) Oligonucleotide Synthesis: A

5   Practical Approach, IRL Press, Oxford; Blackburn et al. (1990) Nucleic Acids in Chemistry and Biology, IRL Press, Oxford, especially Chap. 3; Chaps 13-16 of Methods in Enzymology, Vol. 154 (Wu et al., eds.) Academic Press, San Diego, CA, 1987; Chaps. 13-14 in Methods in Molecular Biology, Vol. 4 (Walker, ed.), Humana

10   Press, Clifton, NJ, 1988; and Uhlmann et al. (1990) Chem. Rev. 90:544-584.

     In addition to providing methods for chemical synthesis of DNA and RNA, some of these references (particularly Gait and Uhlmann et al.) describe the reactions and methodology for adding protecting

15   groups to 5'OH, 3'OH and 2'OH groups, and exocyclic amine groups (N^6 of adenine, N^4 of cytosine, N^2 of guanine). These references also provide information and protocols to attach nucleotides to solid supports which protocols are useful for attaching the base-protected nucleotide analogs of this invention.

20   The synthesis of 5-SDNP-dC and 5-SDNP-C can be accomplished by reaction of 5-Br-dC or 5-Br-C with sodium hydrogen sulfide, as described by Solan, followed by conversion into the DNP thiol ether. In these cases, it may be necessary to first protect the 4 amino group of cytosine (e.g., as the iso-butoxy amide) and

25   complete the remaining protection steps (5'-O-DMTr and 2'-O-TBDMS, if needed) before preparation of a CED phosphoramidite or

-25-

PCT/US97/12888

conversion to the triphosphate (Sung, 1982, J. Org. Chem. 47:3623). Alternatively, 5-S-dU or 5-S-U can be converted to the corresponding compounds. Several methods are available for this conversion of U to C, including Sung, Xu, and MacMillan.

The conversion of 5-SDNP-dU and 5-SDNP-U to 5'phosphates, phosphoramidites, phosphorothioamidites, phosphonates, and O-substituted monophosphates has been described herein above. All of these reaction schemes can be used to produce the corresponding thiol protected cytosine analogs. If necessary, various protecting groups for the 5'OH, 3'OH or 2'OH groups as well as the exocyclic amines can be added in accordance with the methodology described herein.

Another aspect of this invention relates to the oligonucleotides or polynucleotides containing the nucleotide analogs of this invention and a method of preparing such nucleic acids using the subject nucleotide analogs. Oligonucleotides and polynucleotides of this invention are made by standard methods of chemical (automated or manual) synthesis or enzymatic synthesis of DNA and RNA. Such methods are well known in the art. In chemical synthesis, the nucleotide analog of this invention is substituted for a particular nucleotide at the desired point in the synthesis.

After incorporation of the nucleotide analog and complete synthesis of the oligonucleotide or polynucleotide, the thiol can be deprotected and reacted with any number of thiol-modifying reagents to attach a functional group at that point on the oligonucleotide or polynucleotide. Deprotection of S-DNP or S-S-Et

-26-

can be accomplished by treatment with ß-mercaptoethanol or by other means of reducing sulfides. In a preferred method for the DNP analog, deprotection is accomplished by treating the oligonucleotide with 1.4M BME at 45°C for 4 hours or at room temperature overnight.

Any variety of functional groups can then be added to the reactive thiols group generated by the deprotection step. Such functional groups include cross-linking groups, photoactive cross-linking groups (e.g. arylazides), and reporter molecules such as radioisotopes, biotin, enzymes and fluorescent markers. The methods for adding functional groups are well know in the art.

Yet another aspect of this invention provides a method of preparing the nucleotide analogs of the invention. In particular, this method involves preparing a thiol protected nucleoside or nucleotide base wherein said thiol is attached to a position on said base that is not involved in Watson-Crick base pairing or does not disrupt normal Watson-Crick base pairing; reacting said nucleoside or nucleotide base under conditions to effect conversion of said base to a phosphoramidite, phosphorothioamidite, phosphonate, O-substituted monophosphate or phosphate nucleotide analog and under conditions which do not destroy the protected thiol; and recovering said analog. In accordance with this invention, this method is accomplished as described above for synthesis of the subject nucleotide analogs. Recovery of the analogs can be accomplished by HPLC, FPLC or other chromatographic separation techniques.

-27-

A further aspect of this invention provides a method of synthesizing a nucleic acid with a functional group by incorporating a thiol protected nucleotide analog in accordance with this invention into a nucleic acid by a chemical or enzymatic method for nucleic acid synthesis; recovering the nucleic acid containing the analog; deprotecting the analog of that nucleic acid to produce a nucleic acid containing a reactive thiol group; reacting the reactive thiol group with a thiol-modifying reagent to thereby attach a functional group and produce the nucleic acid with the functional group; and recovering the nucleic acid with the functional group. In accordance with this invention, this method is accomplished as described hereinabove for synthesis of oligonucleotides and polynucleotides (See for example Gait or Sambrook et al). As used herein, nucleic acids include oligonucleotides and polynucleotides. Preferably, the oligo-nucleotides range in size from about 5 to about 100 nucleotides. Polynucleotides range in size from 100 nucleotides to 10 kb or more. Recovery of the analogs can be accomplished by HPLC, FPLC, other chromatographic techniques, extraction, phase separation or precipitation.

While the invention will now be described in connection with certain preferred embodiments in the following examples so that aspects thereof may be more fully understood and appreciated, it is not intended to limit the invention to these particular embodiments. On the contrary, it is intended to cover all alternatives, modifications and equivalents as may be included

-28-

within the scope of the invention as defined by the appended claims. Thus the following examples, which include preferred embodiments, will serve to illustrate the practice of this invention, it being understood that the particulars shown are by

5   way of example and for purposes of illustrative discussion of preferred embodiments of the present invention only and are presented in the cause of providing what is believed to be the most useful and readily understood description of formulation procedures as well as of the principles and conceptual aspects of the

10   invention.

EXAMPLES

Procedures

Materials. Unless otherwise stated, starting materials for the chemical synthesis of the phosphoramidite were obtained from

15   Aldrich, Sigma, Fluka or Fisher Scientific and were used without further purification. Anhydrous HOAc was prepared by fractional freezing (Nagamachi et al., 1974). Anhydrous CH_2Cl_2 was distilled from phosphorus pentoxide. Anhydrous CH_3CN was purchased from Cruachem. Dowex 50W x4 was purchased from BioRad. Analytical thin

20   layer chromatography was performed on Whatman silica gel 60 plates with fluorescent indicator. Column chromatography was performed with silica gel 60A from American Scientific Products. Reagents for the oligonucleotide syntheses were obtained from Cruachem. Standard exocyclic amine protecting groups were used: Benzoyl for

25   dA and dC and isobutyryl for dG. Solvents utilized for the reverse

-29-

phase chromatographic purification of the oligonucleotides were or HPLC grade. The fluorescent thiol modifying agent, 5-IAF was purchased from Molecular Probes (Eugene, Oregon).

Buffers. The buffers used were as follows: buffer A, 100 mM
5   TEAA (pH 7.1); buffer B, 20 mM TEAA (pH 7.1); buffer C, 50 mM Tris-HCl (pH 4.9), 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$; buffer D, 50 mM Tris-HCl (pH 9.0), 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 0.5 mM spermidine; buffer E, 7 M urea, 0.05% (w/v) xylene cyanol; buffer F, 5% (v/v) glycerol, 0.04% (w/v) xylene cyanol, 0.04% (w/v) bromophenol blue.

10   Enzymes and Antibodies. Nuclease P1 was obtained from Pharmacia Biotech. CIP was supplied by Promega. T4 Polynucleotide Kinase was purchased from New England Biolabs. AmpliTaq Polymerase was obtained from Perkin Elmer. Rabbit polyclonal anti-DNP antibody was purchased from Sigma. Anti-rabbit Ig, horseradish
15   peroxidase-linked whole antibody was obtained from Amersham.

Analytical Methods. Melting points were determined on a Mel-Temp melting point apparatus and are uncorrected. ^1H NMR and ^{31}P NMR spectra were recorded on a Varian XL-300 or 500 spectrometer, respectively. Chemical shifts are reported in parts per million
20   (δ) relative to internal tetramethylsilane or to external 85% phosphoric acid. UV spectra were measured on either a LKB Biochem Ultraspec II or a Beckman DU 7500 spectrophotometer. IR spectra were recorded with a Perkin Elmer 710B Infrared spectrometer. FAB MS were recorded on a VG analytical ZAB-E spectrometer. Unless
25   otherwise stated, all chemical reactions were performed under a nitrogen atmosphere.

-30-

HPLC analyses were performed with a Beckman LC system which was equipped with a 126 solvent delivery module and a 168 diode array detector and was controlled with Beckman System Gold Software. A reverse phase C18 column (Beckman Ultrasphere ODS, 250 x 4.6 mm, i.d.) and guard column (Upchurch, 1 cm x 4.3 mm, i.d.) were utilized with both analytical and semi-preparative separations. Sample loops of 1 mL and 50 μL were used for semi-preparative and analytical analyses, respectively. All HPLC analyses were performed using a gradient solvent system composed of triethylammonium buffers and acetonitrile at a flow rate of 1 mL/min. Unless otherwise stated, UV absorption was monitored at a wavelength of 260 nm.

Abbreviations: DNP,2,4-dinitrophenyl; 5-IAF, 5-iodoacetamido-fluorescein; CIP, Calf intestinal alkaline phosphatase; TEAA, triethylammonium acetate; TEAB, triethylammonium bicarbonate; DMT, dimethoxytrityl; RP HPLC, reverse phase high pressure liquid chromatography; PCR, polymerase chain reaction; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; DMTCl, dimethoxytrityl chloride; DMT⁺BF⁻₄, dimethoxytrityl tetrafluoroborate; BME, β-mercaptoethanol; HR FAB MS, high resolution fast atom bombardment mass spectroscopy.

Example 1. Preparation of 5-SDNP-dU-phosphoramidite
Shown in Figure 1 and described in detail below is the pathway used for the synthesis of the DNP-labeled phosphoramidite, Compound 7.

-31-

Synthesis   of   3',5'-O-Diacetyl-5-thiocyanato-2'-deoxyuridine (Compound 2).

Anhydrous HOAc (150 mL) was added to a three-neck round bottom flask equipped with a vacuum adapter and teflon needle. Dry Cl_2 5   was slowly bubbled through the teflon needle into the reaction vessel for a brief period. A 5 mL aliquot of the Cl_2/HOAc solution was removed and subjected to iodometric titration to determine the amount of Cl_2 added (2.9 g, 41 mmol) (Flaschka et al., 1969). KSCN, previously dried overnight at 100°C, (4.4 g, 45 mmol) was 10   added to the reaction vessel. The reaction mixture was allowed to stir at room temperature for 30 min. prior to the addition of 3',5'-O-diacetyl-2'-deoxyuridine,3',5'-O-Diacetyl-2'-deoxyuridine was prepared by reaction of deoxyuridine with acetic anhydride in pyridine for an 82% yield: mp=106-108°C(Lin & Gao, 1983) (1.1 g, 15   3.5 mmol). After the mixture was stirred for an additional 1.5 hours, cyclohexene (10 mL) was added. The mixture was then stirred for 15 min. and filtered; the filtrate was concentrated under vacuum. Residual HOAc was removed azeo-tropically with toluene and the residue was triturated with petroleum ether. The petroleum 20   ether was decanted and the residue subjected to column chromatographic purification [petroleum ether: EtOAc(2:1 - 1:3)] which yielded a semi-solid. Recrystallization from EtOH:EtOAc provided Compound 2 as a white solid (1.05 g; 82%):mp 154-156°C; IR(KBr) $2140cm^{-1}$ (SCN stretch); ^{1}H NMR (CDCl$_3$) δ 9.12(bs, 1H, N$_3$H), 25   8.18 (s, 1H, H$_6$), 6.14 (dd, J=7.8, 5.7 Hz, 1H, H$_{1'}$), 5.13 (m, 1H, H$_{3'}$), 4.24 (m, 1H, H$_{5'}$), 4.19-4.15 (m, 2H, H$_{4'}$ and H$_{5'}$), 2.32 (ddd, J=

-32-

14.4, 5.7, 2.1 Hz, 1H, $H_{2'}$), 2.13-2.07 (m, 7H, 2 CH_3 and $H_{2'}$); UV (MeOH) 270 nm.

Synthesis of 3',5'-O-Diacetyl-5-mercapto-2'-deoxyuridine, (Compound 3).

5   This reaction should be performed in a fume hood. EDTA (25 mL; 0.1 M, pH 7.7) and DTT (293 mg, 1.90 mmol) were added to a solution of thiocyanate 2 (207 mg, 0.561 mmol) in MeOH (30 mL). The reaction mixture was allowed to stir at room temperature for 35 min. and was then filtered. The filtrate was reduced to half by

10   rotary evaporation, made acidic with 10% (v/v) H_2SO_4 and extracted with CH_2Cl_2. The CH_2Cl_2 extract was concentrated under reduced pressure and the residue was triturated with H_2O to provide Compound 3 as a white solid (148 mg; 77%). TLC analysis showed a yellow spot upon development with 2,2'-dithiobis(5-nitropyridine)

15   (Grassetti & Murray, 1969): mp=159-163°C; IR(KBr) 2530 cm^{-1} (SH stretch); ^{1}H NMR ($CDCl_3$) δ 8.70 (bs, 1H, N_3H), 7.67 (s, 1H, H_6), 6.30 (dd, J=8.2, 5.8 Hz, 1H, $H_{1'}$), 5.22 (m, 1H, $H_{3'}$), 4.43-4.25 (m, 3H, $H_{4'}$ and $H_{5'}$), 3.60 (s, 1H, SH), 2.50 (ddd, J=14.4, 5.8, 2.2 Hz, 1H, $H_{2'}$), 2.21-2.10 (m, 7H, 2 CH_3 and $H_{2'}$); UV (0.1 M EDTA w/DTT, pH

20   7.76) 260, 334 nm. (The broad melting point range may indicate the presence of 5,5"-dithiobis(3',5'-diacetyldeoxyuridine). On the average, 7% disulfide analog was present in the sample as determined by UV absorbance (λmax 335 nm) in the presence and absence of DTT (Bardos & Kalman, 1966)).

-33-

Synthesis of 3',5'-O-Diacetyl-5-S-(2,4-dinitrophenyl)mercapto-2'-deoxyuridine, (Compound 4).

2,4-Dinitrofluorobenzene (57 mg, 0.31 mmol) and Et_3N (200 μL) were added to a solution of mercaptan 3 (86 mg, 0.25 mmol) in anhydrous CH_3CN (10 mL). The reaction mixture was allowed to stir at room temperature for 1.5 hours before the solvent was removed by rotary evaporation. The crude material was purified using column chromatography [Hexanes: EtOAc(4:1 to 1:1)]. The product was obtained as a yellow solid (113 mg, 89%): IR(KBr) 1530, 1340 cm^{-1} (NO_2 stretch); ^1H NMR ($CDCl_3$) δ 9.14 (d,J=2.4 Hz, 1H, H_3 of DNP), 8.40 (bs, 1H, N_3H), 8.31 (dd, J=9.1, 2.5 Hz, 1H, H_5, of DNP), 8.23 (s, 1H, H_6), 7.33 (d,J=9 Hz, 1H, H_6 of DNP), 6.32 (dd, J=7.8, 5.7 Hz, 1H, H1'), 5.25, (m, 1H, $H_{3'}$), 4.44-4.25 (m, 3H, $H_{4'}$ and $H_{5'}$), 2.55 (m, 1H, $H_{2'}$), 2.27 (m, 1H, $H_{2'}$), 2.13 (s, 3H, CH_3), 2.10 (s, 3H, CH_3): UV (MeOH) 264, 323 nm.

Synthesis of 5-S-(2,4-Dinitrophenyl)mercapto-2'-deoxyuridine, (Compound 5).

NaOMe (25% solution in MeOH; 82 μL; 0.36 mmol) was added to a solution of Compound 4 (92 mg, 0.18 mmol) in anhydrous MeOH (6 mL) and the reaction mixture was maintained at room temperature. After 2 hours, the mixture was treated with Dowex 50W x4 and filtered. The filtrate was reduced by rotary evaporation. The product was obtained as a yellow solid following column chromatographic purification of the reaction mixture [EtOAc: petroleum ether (1:4 to 1:0)] (62 mg; 81%): IR(KBr) 1530, 1340 cm^{-1} (NO_2 stretch); ^1H NMR

-34-

(Cd₃OD) δ 9.08 (d, J = 2.4 Hz, 1H, H₃ of DNP), 8.85 (s, 1H, H₆),
8.39 (dd, J = 9, 2.4 Hz, 1H, H₅ of DNP), 7.64 (d, J = 9 Hz, 1H, H₆
of DNP), 6.33 (t, J = 6.3 Hz, 1H, H₁·), 4.45 (m, 1H, H₃·), 4.00 (m,
1H, H₄·), 3.84 (dd, J - 12, 3 Hz, 1H, H₅·), 3.75 (dd, J = 12, 3 Hz,
5   1H, H₅·), 2.50-2.33 (m, 2H, H₂·); UV (0.1 M TEAA, pH 7.1 with 57%
CH₃CN) 262, 334 nm.

**Synthesis of 5'-O-Dimethoxytrityl-5-S-(2,4-dinitrophenyl)mercapto-
2'-deoxyuridine (Compound 6).**

Anhydrous CH₃CN (6 mL) was added to a mixture of compound 5
10 (42 mg, 0.10 mmol), DMT⁺BF₄⁻(57 mg, 0.14 mmol) and 2,6-di-*tert*-
butyl-4-methylpyridine (DBMP) (35 mg, 0.18 mmol). The resulting
red solution was heated to reflux. After 1.5 hours, a second
portion of DMT⁺BF₄⁻ (26 mg, 0.064 mmol) and DBMP (13 mg, 0.067 mmol)
was added and the reaction mixture was maintained at reflux for an
15 additional 8 hours. The CH₃CN was removed by rotary evaporation.
CH₂Cl₂ and 5% aq. NaHCO₃ were added to the residue. The mixture was
stirred briefly and the resulting two layers separated. The NaHCO₃
layer was re-extracted twice with additional CH₂Cl₂. The organic
layers were combined, dried (Na₂SO₄) and reduced. The product was
20 purified using column chromatography [CH₂Cl₂:MeOH:pyridine
(99.5:0:0.5 to 98.5:2:0.5)] (27 mg: 37%); ¹H NMR (CDCl₃) δ 9.02
(d, J = 2.1 Hz, 1H, H₃ of DNP), 8.46 (s, 1H, H₆), 8.07 (dd, J = 9,
2.4 Hz, 1H, H₅ of DNP), 7.35-7.09 (m, 10H, H6 of DNP and Ar), 6.75-
6.64 (m, 4H, Ar), 6.41 (dd, J = 9, 2.4 Hz, 1H, H₁·), 4.70 (m, 1H,
25 H₃·), 4.16 (m, 1H, H₄), 3.73 (2s, 6H, 2CH₃), 3.38 (d, J = 7.5 Hz, 2H,

-35-

H$_5$,), 2.64 (m, 1H, H$_{2'}$), 2.44 (m, 1H, H$_{2'}$); UV (0.1 M TEAA, pH 7.1 with 82% CH$_3$CN) 264, 324 nm; HR FAB MS calculated for C$_{36}$H$_{32}$N$_4$O$_{11}$S (M)$^+$728.1788 found 728.1802 (1.9 ppm). (Synthesized by the method of Bleasdale et al. (Bleasdale et al., 1990)).

Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-S-(2,4-dinitrophenyl) mercapto-2'-deoxyuridine-3'-O-(2-cyanoethyl-N,N'-diisopropyl) phosphoramidite (Compound 7).

Compound 6 (13 mg, 18 µmol) was dried under vacuum in the reaction vessel. Anhydrous CH$_3$CN (0.3 mL) was added and the reaction vessel was transferred to an atmosphere bag filled with dry N$_2$. A mixture of 2-cyanoethyl-N,N,N'N'-tetraisopropyl-phosphorodiamidite (5.5 µL, 17 µmol) and tetrazole (1.3 mg, 18 µmol) in anhydrous CH$_3$CN (50 µL) was added. The reaction mixture was maintained at room temperature for 30 min. An additional mixture of 2-cyanoethyl-N,N,N'N'-tetraisopropylphosphorodiamidite (2.3 µL, 7.2 µmol) and tetrazole (0.6 mg, 8.3 µmol) in anhydrous CH$_3$CN (25 µL) was added and the reaction mixture was stirred for 30 min. CH$_2$Cl$_2$ (20 mL) and 5% (w/v) aq. NaHCO$_3$ (15 mL) were added to the mixture; the organic layer was removed and re-extracted with water, dried (Na$_2$SO$_4$) and evaporated. The residue was subjected to column chromatographic purification [CH$_2$Cl$_2$:MeOH:pyridine(99.5:0:0.5 to 98.5:1.0:0.5)] to obtain Compound 7 (8.6 mg, 52%). The yellow product was analyzed by analytical RP HPLC using a gradient of buffer A and acetonitrile [70% (v/v) CH$_3$CN to 80% in 3 min.; 80% CH$_3$CN to 100% in 15 min.]. UV absorption was monitored at

wavelengths of 254 and 330 nm. The diastereomers eluted at 11.09 and 12.49 min. UV scans of the diastereomers showed absorption peaks at 264 and 327 nm. Phosphorus NMR indicated the presence of two diastereomers [^{31}P NMR (CDCl$_3$) δ 149.12, 149.27 ppm].

Example 2. Automated Synthesis of an Oligonucleotide Containing 5-SDNP-dU

Synthesis and Purification of Oligonucleotides.

Oligonucleotides were synthesized on an Applied Biosystems 392 DNA Synthesizer using the standard β-cyanoethyl-protected phosphoramidite method. A series of three modified oligonucleotides were prepared on a 40 nmole scale using a Trityl On method. The thiol-protected analog, Compound 7, was manually dissolved into anhydrous acetonitrile at a concentration of 0.027 M and the solution was attached to the spare port on the DNA synthesizer. As the concentration of Compound 7 was lower than the concentration of the commercial phosphoramidites used in these syntheses (0.05 M), the coupling time for the modified analog was increased by 120 seconds. Syntheses of the corresponding unmodified oligonucleotides were also performed. The sequences of the oligonucleotides are shown in Table 1.

The oligonucleotides were removed from the column by treatment with concentrated NH$_4$OH(1 mL). Deprotection of the exocyclic amines was then accomplished by treatment of the oligonucleotides with an additional 1 mL of fresh concentrated NH$_4$OH at room temperature for 44 hours. The NH$_4$OH solution was concentrated

-37-

using a Savant Speed Vac without heat. Small portions of Et$_3$N were periodically added to the concentrating solution to maintain a basic pH. The crude oligonucleotides were re-dissolved in 10 mM TEAB (pH 7.1) and purified as the DMT ethers using semi-preparative RP HPLC. Two different gradients of buffer B and acetonitrile were utilized. The shorter oligonucleotides (7 and 22 nucleotides) were purified with an elution gradient of CH$_3$CN from 5% (v/v) to 30% in 15 min. then 30% to 100% in 5 min. The 24-mer oligonucleotides were purified with an elution gradient of CH$_3$CN from 5% to 30% in 18 min. then 30% to 100% in 5 min. The appropriate fractions were evaporated to dryness using a Jouan Concentrator without heat. Following purification of the oligonucleotides, the DMT group was removed by treatment with 80% (v/v) HOAc (200 μL) at room temperature for 20 min. The oligonucleotides were ethanol precipitated and re-dissolved into 300 μL of nanopure water and stored at -20°C. The purity of the oligonucleotides was analyzed by analytical RP HPLC (Figure 2A, Oligonucleotides A and B shown). Elution of the oligonucleotides was performed using a linear gradient of CH$_3$CN (5 to 20% in 15 min.) with buffer B. The purified oligonucleotides were also analyzed by 25% polyacrylamide-7 M urea gel (acrylamide/methylene bisacrylamide = 19/1) (Figure 3A, Oligonucleotides C-F shown).

Enzymatic Digestion with Nuclease P1 and CIP.

The oligonucleotides A and B were subjected to enzymatic digestion. The pure oligonucleotides were incubated with Nuclease

-38-

P1 (2 U) at 37 °C for 6 hours in a total volume of 50 μL of buffer
C. Five microliters of 10X buffer D and 43 μL of H₂O were added to
the reaction. CIP (2 U) was added and the pH of the reaction was
adjusted with 5% (v/v) NaOH (0.8 μL). The reaction mixture was
incubated overnight at 37°C. To recover the nucleosides, 11 μL of
3 M NaOAC, pH 5.3 and 250 μL of absolute EtOH were added to the
reaction mixture. The mixture was stored at -80°C for 4 hours and
then centrifuged at 14000 rpm for 20 min. The supernatant fraction
was removed and concentrated. The resulting residue was re-
dissolved into 35 μL of buffer A and analyzed by analytical RP HPLC
(Figure 2B). The elution gradient consisted of buffer A with an
increasing concentration of CH₃CN (0% to 5% in 5 min., 5% to 6% in
15 min., 6% to 50% in 5 min. and finally 50% to 80% in 15 min.).
The UV absorption was monitored at 260 and 330 nm. Commercial
nucleosides and monophosphates of dA, dC, dG and T were utilized as
standards. Synthetic 5-S-(2,4-dinitrophenyl) mercaptodeoxyuridine
(Compound 5) was also utilized as a standard. An enzymatic
digestion without oligonucleotide was performed to identify
background peaks arising from the buffer and enzymes.

Deprotection and alkylation of a DNP-modified oligonucleotide.

Oligonucleotide F (100 pmol) was radiolabeled by reaction with
T4 Polynucleotide Kinase using [γ-³²P]ATP. The labeled oligonucleo-
tide F was isolated by ethanol precipitation and re-dissolved into
30 μL H₂O. The oligonucleotide (4.5 μL) was treated with 1.4 M BME
either at 45 °C for 4 hours or at room temperature for 12 hours in

-39-

a 20 μL solution of 20 mM Tris-HCl, pH 8.2. Following this deprotection to remove the DNP group, half of the reaction mixture was precipitated with ethanol while the other half of the mixture was subjected to alkylation conditions with 5-IAF. The corresponding unmodified oligonucleotide E was treated identically to serve as a control.

To recover the deprotected oligonucleotides by ethanol precipitation, 0.36M NaOAc, pH 5.6 (100 μL), 1M MgCl$_2$(1.2 μL) and nanopure water (10 μL) were added to the reaction mixture (10 μL) to provide a final volume of 121 μL. Absolute ethanol (300 μL) was added and the mixture was cooled at -80°C. The oligonucleotides were pelleted by centrifuging at 14000 rpm for 20 min. The supernatant was removed and the pellet was washed with cold 95% (v/v) ethanol. The mixture was chilled for 20 min. in a dry ice/ethanol bath and spun at 14000 rpm for 20 min. Following the removal of the supernatant fraction, the recovered oligonucleotides were dissolved into 20 μL of nanopure water. Mixtures of oligo-nucleotides (4.5 μL), 100 mM DTT (0.5 μL) and buffer E (5 μL) were prepared and analyzed by electrophoresis on a 25% polyacrylamide-7 M urea gel (acrylamide/methylene bisacrylamide = 19/1) (Figure 3B).

Alkylation of the deprotected thiol moiety was accomplished by the addition of 15 μL of 20 mM Tris-HCl (pH 4.8) to 10 μL of the deprotection reaction (pH 8.2). Following this addition, the pH of the alkylation reaction was 7 (as judged by pH paper). One microliter of 10mM 5-IAF in DMF was added and the reaction mixture was maintained at room temperature overnight. The oligonucleotides

-40-

were recovered by ethanol precipitation and re-dissolved into 20 μL nanopure water. The alkylated oligonucleotides (4.5 μL) were mixed with 100 mM DTT (0.5 μL) and buffer E (5 μL) and analyzed by electrophoresis on a 25% polyacrylamide-7M urea gel (acrylamide/methylene bisacrylamide = 19/1) (Figure 3B).

DNA synthesis via PCR using a DNP-modified oligonucleotide.

PCR reactions were performed on a Perkin Elmer Cetus Gene Amp PCR System 9600. The DNA fragment containing the bacteriophage lambda PR' promoter, 6S gene, and tR' terminator, with the lambda qut site (Yang et al., 1987) was amplified from plasmid pHA100 (Zhang & Hanna, 1995) using either oligonucleotide E or F and the universal T7 primer (Stratagene). The DNA product was purified by excision of the band from an agarose gel and extraction of the DNA using the commercially available GENECLEAN kit (BIO 101 Inc., La Jolla, CA). The PCR products were dissolved into equal volumes of buffer F and analyzed by electrophoresis on a 1.5% agarose gel. Detection of DNA was achieved by addition of ethidium bromide to the electrophoresis buffer. (Figure 4A).

Immunodetection of DNP-labeled DNA.

Dot blot experiments were performed manually. The DNP-labeled PCR product and the corresponding unlabeled DNA fragment (1 pmole each) were in 10X SSC solution (1.5 M NaCl, 0.15 M Na₃Citrate, pH7). Samples (13 μL) were applied to a pre-wetted charged nylon membrane (Zeta Probe GT; Biorad). The membranes were air dried,

-41-

rinsed briefly with 2X SSC and heated at 80°C under vacuum for 30
min. The membranes were blocked by overnight incubation at 4°C in
TBS-T (20 mM Tris HCl, 500 mM NaCl, 0.4% (v/v) Tween 20, pH 7.6)
containing 10% (w/v) dried milk. The membranes were then washed

5   with TBS-T two times briefly, one time for 15 min., and two times
for 5 min. The membranes were incubated with a 1:1500 dilution of
anti-DNP antibodies for 1 hour, washed as described above and then
incubated with a 1:10,000 dilution of horseradish peroxidase-linked
whole anti-rabbit Ig for 1 hour. Membranes were then washed one

10 time for 15 min. and 4 times for 5 min. with TBS-T. Detection of
the antibodies was carried out using the ELC chemiluminescence
method (Amersham). The membranes were incubated with the detection
agents for 1 min. before exposure to x-ray film for 2 min. (Figure
4B).

15   RESULTS

The synthesis and characterization of a novel deoxyuridine
phosphoramidite (Compound 7) containing a protected thiol group at
the 5 position of the mercaptopyrimidine ring is described (Figure
1). This DNP-labeled analog was site-specifically incorporated

20 internally into a series of oligonucleotides. Significant
differences in overall synthetic yields were not observed between
oligonucleotides made with normal deoxynucleoside phosphoramidites
and those which contained the analog, indicating that the modified
phosphoramidite is a suitable reagent for automated oligonucleotide

25 synthesis. Enzymatic digestion of the DNP-labeled oligonucleotide

-42-

established the stability of the modified pyrimidine to the conditions for chemical synthesis and purification of oligonucleotides (Figure 2B). Selective deprotection of the 5-thiol moiety with BME and subsequent modification of the unmasked

5   thiol with 5-IAF were verified by gel electrophoresis of the oligonucleotides (Figure 3B). Incorporation of the DNP-labeled oligonucleotide into double-stranded DNA was achieved using PCR. Isolation and characterization of the PCR products included gel electrophoresis and immunodetection with anti-DNP antibodies

10   (Figure 4A and B).

Synthesis.

Synthesis of the deoxyuridine phosphoramidite analog (Compound 7) was carried out following the sequence of reactions outlined in Figure 1. Compounds 2-7 have not been previously described in the

15   literature. The method chosen for incorporation of the thiol group required thiocyanation of the pyrimidine ring. Initially, this reaction was performed on deoxyuridine and the reaction conditions utilized were similar to those previously described (Bradley & Hanna, 1992; Nagamachi et al., 1974; Torrence et al., 1968).

0   Although literature reports indicate yields of 54% for the thiocyanation of deoxyuridine, this was difficult to achieve. The yield for this reaction varied considerably and was at times as low as 9%. A major side product in this reaction was isolated and identified by NMR and IR as 5'-O-acetyl-5-thiocyanatodeoxyuridine

-43-

indicating that along with thiocyanation, acetylation of the sugar ring had occurred.

To circumvent the problems associated with the reaction described above, 3',5'-O-diacetyl-2'-deoxyuridine (Compound 1) was utilized as the starting material. Protection of the sugar hydroxyls by acetylation prior to thiocyanation results in higher yields (95%) for the reaction (Nagamachi et al., 1974). However, removal of the acetyl protecting groups could not be accomplished using standard basic conditions required for deesterification without destruction of the thiocyanato group. Therefore, a synthetic approach was developed which required reduction of the thiocyanato compound to the mercapto analog prior to removal of the acetyl groups (Figure 1). Thiocyanation of the 3',5'-O-diacetyl-2'-deoxyuridine (1) was achieved following the literature procedure described above except that the crude material was extracted with petroleum ether to remove nonpolar impurities prior to column chromatographic purification. The pure product was obtained in 82% yield as a white solid.

Reduction of the thiocyanato derivative 2 to the mercapto analog (Compound 3) was achieved by reaction with DTT in a solvent mixture of MeOH and 0.1 M EDTA (pH=7.8). This reaction is a variation of a reported procedure (Lin et al., 1988) for the reduction of 3'-azido-2',3'-dideoxy-5-thiocyanatouridine; this mercapto analog however was not isolated. Purification of Compound 3 was achieved by acidification of the reaction mixture followed by concentration of the mixture. The residue was extracted with CH_2Cl_2

-44-

and H₂O. The presence of the acetyl groups insured the product was organic solvent soluble. In order to remove residual DTT, the organic layer was reduced and the residue triturated with water to crystallize the product as well as to dissolve residual DTT. Compound 3 was obtained as a white solid in 75% yield.

The mechanism for the reduction of the thiocyanato requires displacement of the cyano group by DTT. Since no attempt was made to remove the cyanide ion before acidification of this reaction mixture, every step of the isolation of Compound 3 from acidification to trituration with H₂O was performed in a fume hood. This synthesis of Compound 3 was designed for the preparation of both 5-alkyldithio and 5-alkylthio or 5-arylthiopyrimidine analogs. To prevent degradation of the 5-alkyldithio compounds by DTT, purification of Compound 3 was necessary. This is not the case with formation of 5-alkylthio pyrimidines (Lin et al., 1988) and this purification step has been eliminated for Compound 4 resulting in a slight decrease in yield.

The DNP protecting group has been employed in oligonucleotide synthesis as a protecting group for 6-mercaptopurine and 6-thioguanine phosphoramidites (Xu et al., 1992a; Xu et al., 1992b) and as a functional moiety on a non-nucleoside phosphoramidite (Grzybowski et al., 1993). These studies indicated the DNP group was stable during automated DNA chemical synthesis. Conversion of mercaptan 3 into the 2,4-DNP analog 4 was achieved by reaction with 2,4-dinitrofluorobenzene in anhydrous CH₃CN with Et₃N present as a catalyst (Xu et al., 1992b). The product was isolated as a yellow

-45-

solid in 88% yield. Attachment of the thiol protecting group prior to removal of the acetyl moieties insured that the product was soluble in organic solvents, which aided in the purification of the material. More importantly, protection of the sulfur inhibits the

5   oxidation of the analog to the disulfide compound, a potential reaction of 5-mercaptopyrimidines (Kalman & Bardos, 1967).

Another method of preparing 5-arylthiol ethers of deoxyuridine has been previously described. Bergstrom et al. has prepared a series of alkyl and aryl mercaptodeoxyuridine analogs using a

10 palladium-mediated reaction between 5-(chloromercuri)-2'-deoxyuridine and the appropriate disulfide (Bergstrom et al., 1991). Although this method offers a simple route to arylthiol ether analogs it was not selected for study. Development of a more versatile intermediate such as Compound 3 which can be easily

15 converted into various types of analogs including disulfides was desired for future studies. Additionally, the synthetic utility of the method has not been fully evaluated and therefore formation of the desired dinitrophenyl analog (Compound 5) was uncertain.

Deesterification of nucleosides is commonly performed under

20 basic conditions using NH_3 or NaOMe (Kini et al., 1989; Trivedi & Bruns, 1989). Literature reports suggest the DNP group is marginally stable to nucleophiles such as NaOMe; the stability of this group being strongly dependent on the reaction conditions (Greene, 1981). Successful removal of the acetyl groups was

25 achieved by reaction of the diacetyl analog (Compound 4) with NaOMe in MeOH without an apparent loss or modification of the DNP group

-46-

as judged by NMR and IR. Following the deacetylation, the diol (Compound 5) was isolated as a yellow solid in 81% yield. An initial attempt to deacetylate Compound 4 under mild reductive methods was performed (Brown et al., 1982; Soai & Ookawa, 1986). However, reaction with LiBH$_4$ in a mixture of Et$_2$O and MeOH did not result in the formation of Compound 5, possibly due to the poor solubility of the starting material.

Incorporation of the DMT protecting group at the 5′ position of nucleosides is typically accomplished by reaction with DMTCl in the presence of pyridine or dimethylpyridine (Chaudhary & Hernandez, 1979). However, as utilization of these conditions with other 5-thiol modified nucleosides was not successful (unpublished results), this reaction was not attempted for the synthesis of Compound 6. Conversion to the 5′-dimethoxytrityl analog (Compound 6) was achieved by reaction of Compound 5 with DMT$^+$BF$_4$− in the presence of DBMP in CH$_3$CN at reflux. This reagent, DMT$^+$BF$^-_4$, has only recently been utilized in the preparation of 5′ protected nucleosides (Bleasdale et al., 1990; Lakshman et al., 1992).

The proposed mechanism for dimethoxytritylation with DMTCl and pyridine requires formation of a dimethoxytrityl pyridinium salt which then undergoes reaction with the 5′ hydroxyl group (Chaudhary & Hernandez, 1979). Formation of the pyridinium salt is considered the limiting step. The reagent DMT$^+$BF$^-_4$ serves as a source for the dimethoxytrityl cation. By preforming the cation prior to reaction, the slow step in the dimethoxytritylation has been eliminated and the ether should be more easily formed. Although

-47-

Compound 6 was obtained using this reagent, the reaction was performed under reflux conditions. It is unclear why the reaction requires such stringent conditions but presumably steric hindrance plays some part. Modification of other bases has been reported to inhibit dimethoxytritylation possibly due to steric hindrance (Cosstick & Douglas, 1991).

Silica gel column chromatography was employed to isolate Compound 6. Due to the acidic nature of the silica gel, a small percentage of pyridine was added to the eluting solvent to prevent loss of the acid sensitive DMT group. Initially, Et_3N was used in the solvent system; however triethylammonium salts of Compound 6 were isolated as determined by NMR. This was not observed during the synthesis of the standard 5'-dimethoxytrityl-2'-deoxyuridine. Switching to pyridine, a weaker base, in the eluting solvent system eliminated this problem. A satisfactory high resolution FAB mass spectra was obtained for Compound 6.

The Compound 6 was isolated as a yellow solid in yields ranging from 35 to 48%, making formation of the dimethoxytrityl ether the lowest yielding reaction in this synthetic sequence. However, approximately 20% of the starting material, Compound 5, was converted into the 3',5'-bis(dimethoxytrityl) analog, which was isolated and converted back into 4 for recycling into the synthesis.

The final reaction in the preparation of Compound 7 is the incorporation of the phosphoramidite moiety at the 3' position of the nucleoside. This is commonly accomplished by reaction of the

-48-

appropriately protected nucleoside with 2-cyanoethyl-N,N-diisopropyl chlorophosphoramidite in the presence of the N,N-diisopropylethylamine (DIEPA) (Sinha et al., 1984). Utilization of these conditions did not result in the desired product and starting
5   material was re-isolated. However, a standard reaction with 5'-dimethoxytrityldeoxyuridine was successful. A second method for incorporation of the phosphoramidite moiety which utilizes the reagent 2-cyanoethyl-N,N,N,'N'-tetraisopropylphosphorodiamidite in the presence of tetrazole in CH$_3$CN was evaluated (Barone et al.,
10   1984). Unlike the basic reaction with DIEPA, these slightly acidic conditions provided the DNP-labeled phosphoramidite, Compound 7. RP HPLC analysis of Compound 7 indicated the presence of a pair of closely eluting peaks (11.0 and 12.4 min) representative of the expected 1:1 mixture of diastereomers. The overall purity of the
15   sample was determined to be greater than 96% and no starting material was observed in the chromatogram. HPLC data was quantified using a simple area percent method. Similar results were observed using silica gel TLC analysis (CH$_2$Cl$_2$:MeOH:Et$_3$N;96:3:1). The phosphorous NMR spectrum showed two
20   characteristic peaks associated with the diastereomeric phosphoramidites (149.12 and 149.27 ppm). An additional peak appeared in the NMR at 14.3 ppm. Based on the chemical shift of this peak, the impurity is presumed to be a hydrolysis product (Scremin et al., 1994). No attempt was made to remove this
25   material prior to oligonucleotide synthesis as it is not reactive in the oligonucleotide chemistry. A high resolution FAB mass

-49-

spectrum could not be obtained for Compound 7 due to its
instability during analysis. This was not surprising as the
standard deoxyuridine phosphoramidite displayed a high degree of
instability during high resolution FAB MS analysis; however, a
5   small molecular ion peak was identified in the spectrum.
Evaluation of Compound 7 using low resolution mass spectroscopy
showed a fragment (m/z=407) consistent with the loss of the DMT and
the phosphoramidite moieties. As a similar fragmentation pattern
was observed for the standard deoxyuridine phosphoramidite under
10   these experimental conditions, the information supports the
proposed structure.

Syntheses and purification of oligonucleotides.

Internal incorporation of the phosphoramidite analog, Compound
7, into series of oligonucleotides has been accomplished using an
15   automated DNA synthesizer. The natural oligonucleotides with
corresponding sequences were prepared as standards. The sequences
for the oligonucleotides which are listed in Table 1 were chosen
for two reasons: the 7-mer was designed for characterization
purposes while the 22-mer and the 24-mer were developed to study
20   the DNA-protein interactions during transcription of the
bacteriophage lambda 6S gene. The concentration of the
phosphoramidite solution of Compound 7 was diluted twofold as
compared to the normal phosphoramidites (dA, dG, T and dC) utilized
in these DNA syntheses. This dilution was performed to conserve
25   material. The coupling time for the reaction involving the

-50-

modified phosphoramidite was increased by 120 seconds. Under these reaction conditions, the coupling yield for the modified phosphoramidite was >96% which was comparable to the coupling yields of the other standard phosphoramidites. The average stepwise yields for all six oligonucleotide syntheses were determined automatically by the ABI synthesizer using a trityl cation assay. These yields are listed in Table 1.

Selective removal of the exocyclic amine protecting groups in the presence of the DNP moiety was achieved by treatment with NH$_4$OH at room temperature for 44 hours. Typically, this deprotection is performed at elevated temperatures of 55°C for 8 hours (ABI technical bulletin #13, 1987). However, studies involving reaction of the DNP-labeled nucleoside 5 with conc NH$_4$OH at 55°C indicated the analog was not stable to these conditions but was stable at room temperature (data not shown). As successful removal of standard exocyclic amine protecting groups at room temperature has been reported for other oligonucleotides containing modified phosphoramidites, these conditions were utilized (Sheardy & Seeman, 1986). Although these initial studies were performed with the standard exocyclic amine protecting groups, phosphoramidites containing a wide variety of protecting groups which can be efficiently removed under mild conditions are now commercially available. These will be investigated for compatibility with phosphoramidite Compound 7.

The crude oligonucleotides were purified as the DMT ethers by RP HPLC. Removal of the DMT group under standard acidic conditions

-51-

did not affect the DNP moiety. Following removal of the DMT group, the oligonucleotides were analyzed for purity by both analytical RP HPLC and gel electrophoresis as shown in Figures 2A and 3A respectively. The oligonucleotides showed one major product in

5   both analyses. The retention times of the DNP-labeled oligonucleotides were greater than the corresponding natural oligonucleotides in the HPLC analyses due to the lipophilic nature of the dinitrophenyl group. The RP HPLC retention times for the purified oligonucleotides are listed in Table 1. The overall

10   yields of the pure material which were calculated by dividing the optical density of pure oligonucleotide by the optical density of crude DMT-labeled oligonucleotides times 100% are found in Table 1 (Kuimelis & Nambiar, 1994). The yields for the corresponding normal and DNP-labeled oligonucleotides were comparable.

15   **Enzymatic digestion of the heptanucleotide.**

To establish that the DNP group is stable to the conditions of automated DNA synthesis, deprotection, and purification, a DNP-labeled oligonucleotide was enzymatically digested into its corresponding nucleosides and evaluated by RP HPLC. In order to

20   observe the single DNP-labeled nucleoside present in the enzymatic digest, the short chain oligonucleotide B was used for this analysis. The sequence chosen for this heptanucleotide did not contain deoxycytidine. Deoxycytidine co-elutes with 5-mercapto-deoxyuridine, a potential side product arising from the loss of the

25   DNP group, in the gradient system utilized in the analysis.

-52-

Deoxycytidine was therefore omitted from the sequence to eliminate uncertainty in identifying the nucleoside peaks.

5-Mercaptodeoxyuridine was synthetically prepared in this lab by reduction of 5-thiocyanatodeoxyuridine (Nagamachi et al., 1974).

5   Initial degradation of the oligonucleotide B and the corresponding normal oligonucleotide A was attempted by concurrent digestion with snake venom phosphodiesterase (SVP) and CIP (Kuimelis & Nambiar, 1994). This method proved useful in the digestion of the natural oligonucleotide but did not appear to

10   fully degrade the DNP-labeled material. This is consistent with other studies in which SVP has failed to function with some modified oligonucleotides and nucleotides (Gao et al., 1992; Hanna et al, 1989; Hanna et al., 1993). Substitution of the exonuclease SVP with an endonuclease Nuclease P1 resulted in complete

15   digestion. Due to a difference in the pH requirements of these enzymes for activity, this digestion was accomplished in two consecutive enzymatic reactions. First, Nuclease P1 cleaved the phosphodiester bonds to provide the 5'-monophosphates which were then subjected to treatment with CIP. The resulting nucleosides

20   were isolated and analyzed by RP HPLC. The results are shown in Figure 2B.

        This enzymatic degradation provided the correct ratios of unmodified and modified nucleotides. For oligonucleotide A, the ratio of dA, dG and T was 1.9:2.1:2.9. For oligonucleotide B, the

25   nucleoside 5 is also expected in the reaction mixture providing a ratio of dA, dG, T and 5 of 1.9:2.1:1.9:1.1. Chemically

-53-

synthesized Compound 5 was utilized as a standard. Both the enzymatically obtained and the chemically obtained nucleoside Compound 5 eluted with the same retention time during RP HPLC analysis. In addition, these compounds provided identical UV spectra with absorption peaks appearing at 262 and 334 nm. Finally, no unidentifiable peaks were observed in the HPLC chromatograms of the enzymatic digests and only one peak representing nucleoside Compound 5 appeared in the HPLC chromatogram monitored at 330 nm. In this analysis, absorption at 330 nm is associated with the presence of either the dinitro-phenylthiol ether or the 5-mercaptopyrimidine moiety. If the DNP-labeled analog had undergone modification during automated DNA synthesis, deprotection, or purification, extraneous peaks were predicted to appear in this chromatogram.

Deprotection and alkylation of a DNP-modified oligonucleotide.

Removal of the DNP moiety was accomplished by reaction of oligonucleotide F with 1.4 M BME in Tris buffer pH 8 (Shaltiel, 1967). The reaction was performed under two different reaction conditions: room temperature for 12 hours and 45°C for 4 hours. Oligonucleotide E was reacted under the same conditions to serve as a standard. No attempt was made to isolate the deprotected oligonucleotide prior to modification with 5-IAF. The pKa of BME is 9.6 while the pKa of 5-mercaptodeoxyuridine is 5.0 (Bardos & Kalman, 1966). Modification of a thiol by reaction with haloacetamido compounds such as 5-IAF requires formation of the

-54-

thiolate anion. By lowering the pH of the reaction mixture, the selectivity for modification of the oligonucleotide is enhanced over that of BME, a potential side reaction. The oligonucleotide products in both the deprotection and modification reactions were isolated by ethanol precipitation and analyzed by electrophoresis using a 25% polyacrylamide-7 M urea gel. The results are shown in Figure 3B. Lanes 1 through 4 represent reactions utilizing oligonucleotide E. No difference is observed in these four lanes indicating the natural oligonucleotide structure is stable to the conditions needed for deprotection and alkylation of the thiol. Lanes 5 and 7 represent reactions of oligonucleotide F with BME at room temperature and 45°C respectively, while lanes 6 and 8 depict the corresponding alkylation reactions which were carried out at room temperature. No apparent mobility shift is observed between the natural oligonucleotide (lanes 1-4) and lanes 5 and 7, which is consistent with our observations that a 20-mer RNA containing 5-SH-UTP co-migrates in this gel system with 20-mer containing no analog (He et al., 1995). In contrast, a substantial difference is observed following modification with fluorescein, consistent with other reports of gel mobility shifts of fluorescein-labeled oligonucleotides (Clegg et al., 1992; Mergny et al., 1994). In fact, the selection of 5-IAF for the alkylation experiments instead of a thiol-modifying photocrosslinking reagent such as p-azidophenacyl bromide (APB) was due in part to the ease in which the reaction results could be monitored by gel electrophoresis. The conditions for alkylation of the oligonucleotide with 5-IAF

-55-

however should prove useful with other haloacetyl derivatives such
as APB. We have specifically labeled the thiol group in an RNA
molecule containing 5-SH-UMP (He et al., 1995) with APB. We have
also attached APB to a DNA molecule generated by PCR with a
5   oligonucleotide containing (compound 7), after BME removal of the
DNP group (not shown).

As the mobility shift associated with the DNP label is small
(Figure 3A), it is possible that a mixture of DNP-labeled and 5-
mercapto-labeled oligonucleotides may appear as one band on a 25%
10 polyacrylamide-7 M urea gel. Therefore, one can not determine the
extend to which deprotection has occurred. However, the appearance
of a new band in lanes 6 and 8 (Figure 3B) with the subsequent
disappearance of the band seen in lanes 5 and 7 indicates that the
oligonucleotide underwent deprotection and alkylation. Although
15 deprotection does occur with time at room temperature, the reduced
time required to deprotect at 45°C makes this the preferred method.

PCR reaction using a DNP-modified oligonucleotide.

The DNP-labeled oligonucleotide was incorporated into double
stranded DNA using PCR. Plasmid pKA100 served as the template for
20 this reaction. A 492 base pair segment of this plasmid was
amplified using the universal T7 primer and either oligonucleotide
E or F. The products isolated from the reaction were analyzed on
a 10% polyacrylamide-7 M urea gel (data not shown) and a 1.5%
agarose gel (Figure 4A). Differences in mobility were not observed

-56-

WO 98/03532   PCT/US97/12888

between the standard PCR product and the DNP-labeled PCR product in these gel analyses and no difference in yield was observed.

Immunodetection of DNP-labeled oligonucleotides and DNA.

To insure the stability of the DNP protecting group during the PCR reaction, the products described above were subjected to immunodetection studies using rabbit polyclonal anti-DNP antibodies. The natural PCR product served as a negative control. The DNP group was detected using a chemiluminescence method and only the DNP-labeled nucleic acid provided a signal (Figure 4B).

Example 3:   Analysis of Protein-DNA Interactions
Utilizing 5-S-DNP-dU

Photochemical crosslinking is a powerful technique for characterization of both RNA-protein and DNA-protein interactions in nucleoprotein complexes. Photochemical crosslinking "traps" weak or transient associations which might not survive isolation procedures such as immunoprecipitation, gel filtration, or filter binding. There are a number of approaches which can be taken in photochemical crosslinking. One involves direct irradiation of a nucleoprotein complex with short wavelength ultraviolet light. This method relies upon the direct excitation of nucleotides or amino acids to generate chemically reactive species. Such nonspecific labeling can reveal whether a protein-nucleic acid interaction exists, but there are several problems associated with this approach. These include low crosslinking yields, the

-57-

BNSDOCID: <WO___9803532A1_I_>

inability to incorporate a crosslinking group at a specific site in the nucleic acid, and degradation of some proteins during irradiation. A second approach involves the use of nucleotide analogs modified with photoreactive crosslinking groups (reviewed

5   in M. Hanna, 1989, Methods in Enzymology). These groups are chemically inert in the absence of light but are converted to chemically reactive species upon irradiation. Often this can be achieved with long wavelength ultraviolet or visible light, resulting in less nicking of protein or nucleic acid. Molecular

10   contacts can be identified at the level of specific nucleotides in nucleic acids and amino acids in proteins. One can therefore gain structural information about macromolecular complexes which cannot be obtained by most other biochemical approaches.

For such mechanistic studies a variety of nucleotide analogs

15   have been developed and utilized (M. Hanna 1989, Methods in Enzymology: Bradley and Hanna, 1992). These analogs are either tagged with a photoreactive crosslinking group or contain functional groups that can be tagged with crosslinkers after incorporation into DNA or RNA. These analogs can be incorporated

20   into nucleic acids, either enzymatically or chemically, to analyze molecular interactions in protein-nucleic acid complexes. Upon photoactivation, analog-tagged nuecleic acids become covalently attached to adjacent macromolecules (protein, DNA, RNA) with which they have direct interactions. Therefore, nucleic acid binding

25   domains in complexes can be characterized, nucleic acid binding proteins in extracts can be identified, and determinants involved

-58-

in these specific interactions can be characterized at the level of individual nucleotides and amino acids. This methodology allows mechanistic questions about the way that specific protein-DNA or protein-RNA interactions regulate gene expression to be asked.

5   Herein described are the approaches used to characterize the protein-DNA interactions in the E. coli transcription complex which forms on the bacteriophage lambda P_R, promoter.

Preparation of Site-Specifically Modified DNA Oligonucleotides

The method described here is used to probe for interactions

10   between a specific nucleotide in DNA and a DNA-binding protein. In this case, interactions between a specific nucleotide on the non-template strand of the P_R, promoter are sought. The chemical approach to preparation of site-specifically modified DNA involves incorporation of a modified nucleotide during the automated

15   synthesis of an oligonucleotide. Because the conditions used in automated oligonucleotide synthesis are rather harsh, the direct incorporation of a photocrosslinking azide group by this method is not possible, and a means for the post-synthetic incorporation of the azide group is required. 5-S-DNP-dU was incorporated into a

20   DNA oligonucleotide, as described herein. For analysis of protein interactions with the -12 position in the nontemplate strand of the P_R, promoter, an oligonucleotide with the sequence 5'-5D-DNP-dU-AA ATT TGA CTC AAC GAT GGG was synthesized on an Applied Biosystems 392 DNA Synthesizer using the standard β-cyanoethyl-protected

25   phosphoramidite method by attaching the analog to the fifth

-59-

substrate port. After synthesis, removal of the oligonucleotide from the column, and removal of the exocyclic amine protecting groups, the oligonucleotide was ethanol precipitated, dissolved in water and stored at -20°C.

5   **Analysis of 5-S-DNP-dU Effects on Hybridization Properties of ODNs**

To evaluate the effect of incorporation of one or more 5-SDNP-dU analogs into oligodeoxynucleotides (ODNs) on their hybridization properties, the ODNs shown in Table 2 were synthesized. The observed melting temperatures (T_m) showed that incorporation of one

10   DNP analog decreased the melting temperature by only 1.8°C (hybrid V), compared to the normal DNA hybrid (I), and substitution with two DNP analogs caused a decrease of only 4.4°C (hybrid VI). Therefore, the effect of a single substitution is less than that caused by a single mismatch involving unmodified nucleotides

15   (hybrids II, III and IV vs I), and the effect of two modified analogs is still less than that of a single T-C mismatch (hybrid IV). Mismatches of 2 or more are commonly used for site-directed mutagenesis, which involves hybridization of the oligonucleotides to single-stranded DNA. Substitution with the DNP analog should

20   therefore work as well, if not better. The use of ODNs containing one or more 5-SDNP-dU analog in assays requiring specific hybridization to a complementary strand is therefore quite feasible.

By comparison, oligodeoxynucleotides (30-mers) containing a

25   previously described nucleobase fluorescent derivative, 5-Amido-

-60-

(Carboxyfluorescein)-2'-deoxyuridine, caused decreases in melting temperature of 2°C for ODNs with one nucleobase substitution, 6°C for ODNs with two nucleobase substitutions, and 12°C for ODNs containing three nucleobase substitutions (Jadhav et al. 1997). A more serious effect on the hybridization properties was observed for another previously described analog, 2-thiodeoxyuracil. Placement of even a single 2-thiodeoxyuracil analog into oligodeoxynucleotides (14-mers) caused a decrease in T_m of 3-4°C (Kuimelis and Nambiar, 1994).

Table 2. T_m values for native and modified ODNs

ODN 1	5'	GAC	VCA	ACG	AWG	GGX	TAA	TYC	G	3'
ODN 2	3'	CTG	AGT	TAC	TZC	CCA	ATT	AAG	C	5'

ODN	1	1	1	1	2		
Hybrid	V	W	X	Y	Z	$T_m(°C)$	Drop in T_m^a
I	T	T	T	T	A	67.3	0
II	T	T	T	T	G	64.5	2.6
III	T	T	T	T	T	63.6	3.5
IV	T	T	T	T	C	61.5	5.6
V	T	5^b	T	T	A	65.5	1.8
VI	T	5	5	T	A	62.9	4.4
VII	T	5	T	5	A	62.5	4.8
VIII	5	5	T	5	A	58.0	9.3

a - as compared to unmodified duplex (Hybrid I)
b - 5 refers to 5-S-DNP-dU

Attachment of a Crosslinking Group at a Specific Position in DNA

The analog-modified oligonucleotide (100 pmol) was radiolabeled with T4 polynucleotide kinase using [$\gamma^{32}P$] ATP (Figure 5, Step 1), isolated by ethanol precipitation and re-dissolved in water. The presence of the 5-S-DNP group on the 5' terminal

nucleotide did not interfere with recognition and modification of
the 5' OH group by kinase.   The oligonucleotide was then
incorporated into double-stranded DNA by either PCR (not shown) or
by primer-extension.   For primer-extension using a single-stranded
phagemid DNA template, the oligonucleotide (20 pmoles) was annealed
to the template (1 pmol) by heating the mixture to 75°C, cooling to
65°C and then placing the mixture on ice (Step 2).   The
oligonucleotide was then extended with a thermostable DNA
polymerase (AmpliTaq, Perkin-Elmer) by incubating for 5 minutes at
room temperature and then two hours at 70°C.   The template was
ligated with T4 DNA ligase for 1 hour at 37°C, and then ligated DNA
was isolated by ethanol precipitation (Step 3).   Before addition of
a crosslinking group, the DNA protecting group must be removed
(Step 4).   This was accomplished by treatment of the
oligonucleotide (4.5 μL) with 1.4 M β-mercaptoethanol either at
45°C for 4 hours or at room temperature for 12 hours in 20 μL 20 mM
Tris-HCl, pH 8.2.   The deprotected DNA was purified to remove the
β-mercaptoethanol by ethanol precipitation, and the DNA was
dissolved in 20 μL water.   A photocrosslinking group was added to
the DNA by reaction with the alkylating agent azidophenacyl
bromide, and the DNA was again purified for transcription and
crosslinking (Step 5).

Photocrosslinking

For UV crosslinking, the reactions are split in half; a
control aliquot is kept in the dark at room temperature while the

-63-

other aliquot is irradiated at room temperature for 2 minutes in a polystyrene tube 1.5 cm from a 302 nm light source (Spectroline model XX-15B, 1800 μW/cm^2 at 15 cm). After irradiation, DTT (to 60 mM) is added to all samples, and they are left in the dark for at least 5 minutes. After irradiation, but before protein gel electrophoresis, aliquots from the irradiated chase reactions are treated with nuclease to digest the DNA to a small radioactive fragment (Step 6). The nuclease treatment results in digestion of the DNA to small radioactive fragments. This leaves a small radioactive piece of DNA on proteins which have been crosslinked to the DNA during irradiation. The control and irradiated samples are then analyzed by gel electrophoresis and autoradiography.

Electrophoretic Analysis of Crosslinked Proteins

For identification of crosslinked proteins, samples are mixed with an equal volume of 60 mM TrisCl (pH 8.0), 60 mM DTT, 3.4% (w/v) SDS, 17% (v/v) glycerol, 0.02% (w/v) bromophenol blue, 0.02% (w/v) xylene cyanol, heated for 3 minutes at 94°C, and then analyzed on 12 cm x 0.75 cm 10% SDS/polyacrylamide gels with a 4% stacking gel (acrylamide/methylene bisacrylamide = 27/1). Proteins are electrophoretically blotted onto MSI NitroBind Nitrocellulose membrane (Westboro, MA) in 25 mM Tris, 192 mM glycine, 20% (v/v) methanol (Figure 6A). The membrane is then silver-stained in 2% (w/v) sodium citrate, 0.8% (w/v) ferrous sulfate, 0.1% (w/v) silver nitrate for 5 minutes and dried. Autoradiography of the membranes

is at -80°C with Amersham Hyperfilm-MP and a Cronex Lightning Plus intensifying screen.

5-S-DNP-dU can be used for site-specific modification of DNA at internal and terminal positions with DNA, and molecular probes can be placed at variable distances from the DNA backbone. In addition, the commercial availability of antibodies to the DNP group allows the direct use of the DNP-labeled oligonucleotides as immunodetection probes.[15]

Example 4:   Utility: Assays for Diagnosis of Bladder Cancer

Bladder Cancer strikes over 50,000 Americans every year, killing more than 11,000 (23). Currently used means to diagnose bladder cancer usually rely on immunocytochemistry, nuclear DNA content and cytology. Extensive use of highly trained labor that these tests consume make them fairly expensive to administer. In addition, while being fairly specific for the diagnosis of bladder cancer, these tests lack sensitivity. Furthermore, they are poor prognosticators of cancer progression. Tests or biopsies with better sensitivity and prognostic value are invasive, even more expensive and poorly tolerated by patients. These facts result in many patients being under-diagnosed or under-treated for their illnesses.

An ideal test would be sensitive, specific and prognostic for bladder cancer, less invasive, require fewer highly skilled scientists to administer, better tolerated by patients and less costly. Such an assay would be expected to decrease mortality and

-65-

morbidity due to this disease. Such a test would also be expected to represent a significant commercial possibility since it would not only replace many of the existing tests in the marketplace but would enable an increase in the number of tests performed.

5   One of the methods currently used for detection of mRNA levels in a cell utilizes an RNase protection assay (Hershey and krause, 1989) which is commercially available. In an RNase protection assay, a labeled RNA probe of discrete length is constructed that includes two domains (Figure 6). One domain compliments, and

10   therefore hybridizes to, an mRNA species whose abundance one wishes to detect and quantify. This domain forms a double stranded RNA molecule with the intended mRNA target. The other domain is made up of sequence that does not hybridize to the intended target. In an RNase protection assay, total cell RNA (or total cellular mRNA)

15   is incubated with an excess of an RNA probe under conditions that favor specific hybridization of the probe with the mRNA species to be detected and quantified. After the hybridization step, the RNA is digested with a mixture of RNases. All single stranded RNA is hydrolyzed down to nucleotides or short oligonucleotides. Only

20   double stranded RNA is protected from this digestion. All of the unhybridized probe is degraded. This includes the domains of the hybridizing probes that do not complement their targets.

     The digested RNAs are separated according to size by gel electrophoresis and then visualized by autoradiography. A positive

25   signal is a labeled RNA molecule that is the correct length to be that portion of the probe that had hybridized to the intended

-66-

target. This shortened form of the original probe can not be created by other means. This is why false positives are not a technical concern. Also, with an excess of probe and under the appropriate hybridization conditions, nearly all of the intended

5   target is hybridized to the nucleic acid probe (NAP). This makes the intensity of the signal proportional to the original concentration of the target mRNA species.

The advantages of an RNase protection assay are that it is quantitative, reproducible and completely free of false positives

10 derived from technical difficulties. This last feature makes such an assay format appealing as a technological platform for diagnostic determinations of patient derived materials. The disadvantages of an RNase protection assay are that the probes are very tedious to make. They possess very short half lives because

15 of the radioactive isotopes (^{32}P) used to produce the signal. In addition, RNA is inherently less stable than DNA, and some RNases can be extremely difficult to inactivate.

The disadvantages of RNase protection are overcome by the envisioned NAP Nuclease Protection Assays (NAP-protection assays)

20 employing the derivatizable nucleotide analogs herein. The NAPs can be synthesized in large quantities, carefully evaluated (quality control assays), and then stored for future use. There is an economy of scale that this affords which overcomes the difficulty of probe construction and synthesis. These NAPs will

25 utilize non-isotopic means of detection, permitting the NAPs to have extended shelf lives. Additionally, the NAPs for different

-67-

target mRNAs can be modified with different reporter groups. This
will permit the examination of the expression of several genes
simultaneously. This assay can be modified to accept automation to
reduce run times.

5   Use of oligonucleotides or NAP probes which contain modified
nucleotide analogs has been previously reported (21-23, 26-29).
The chemical approach to preparation of site-specifically modified
DNA involves incorporation of a modified nucleotide during the
automated synthesis of an oligonucleotide. The modified

0   oligonucleotide can then be incorporated into double-stranded DNA
by PCR amplification or primer extension of a single-stranded DNA.
The use of various modified phosphoramidites as masked synthons for
preparing derivatized oligonucleotides at both terminal and
internal sites has been reported (Xu, Zheng and Swann, 1992;

15   Maurizi and Ginsburg, 1986). These masked synthons allow
incorporation of protected functional groups such as amines
(Nagamachi, et al. 1974), carboxylic acids, thiols (Goodwin and
Glick, 1993) and thiocarbonyls. Once incorporated the analogs are
deprotected and modified post-synthetically. Convertible

20   nucleoside phosphoramidites, mononomers containing leaving groups,
have also been used to incorporate crosslinking and
photocrosslinking groups by post-synthetic substitution of the
leaving group. Several companies sell reagents for the addition of
molecular tags to either of the two ends of a nucleic acid (5' or

25   3'). Some provide reagents which allow the incorporation of more
than one analog at internal positions in the NAP, to allow an

-68-

increase in the amount of signal produced by the NAP, there is a
resultant perturbation of nucleic acid structure or decrease in
hybridization function (interaction with mRNA) of the NAP. Since
the assay is based on specific NAP hybridization, with the current

5   methods available, increasing the signal in a NAP must be balanced
against the loss of hybridization capacity. However, the goal is
always to obtain the strongest signal possible, because this not
only decreases data acquisition time, but increases the level of
sensitivity of a given assay. This is clearly an important

.0 methodology, for which a commercial market already exists.

Preferred versions of the nucleotide analogs contemplated
herein have the following characteristics which make them
preferable to other currently available nucleotide analogs for this
assay:

15   1.   The nucleotide modification is on a position of the base
               which is not required for normal RNA and DNA basepairing,
               therefore specific hybridization is still possible.

          2.   The reactive group added to the nucleotide is not
               normally found in DNA or mRNA (S group), therefore

20   reaction of the probe modified oligonucleotides or NAPs
               with thiol-reactive alkylating agents gives virtually
               thiol-specific modification of the NAP. Analogs with
               reactive amino groups, which are modified by alkylating
               agents targeting amines, are less specifically modified.

25   This is due to the presence of exocyclic amine groups in

-69-

the normal nucleotides of both DNA and RNA, and these can be modified by the alkylating reagents as well.

3. The reactivity of the aryl thiol group on 5-SH-UTP to alkylating agents is considerably higher than that of alkyl thio ether derivatives ($pK_A=5.5$ vs $8,$).

4. The linker attached to the nucleotide base contains all single-bonds (vs alkyl or alkynyl), thereby allowing free rotation of the probe in the region close to the nucleic acid backbone. This is advantageous because the flexibility of the linker are allows the nucleotide analogs to adopt to different enzyme active sites (in DNA and RNA polymerases), making them generally good substrates for enzymatic incorporation into nucleic acids.

5. As assessed by several functional assays, analogs modified through the aryl thio group of 5-S-U derivatives do not disturb normal RNA secondary or tertiary structures.

Unlike other types of cancer, the primary tumor in bladder cancer sheds or exfoliates relatively large numbers of cells into the lumen of the bladder. These cells can be easily collected in urine or from bladder washes. In a void urine from a patient with a bladder tumor, there will be very few to about a million cancer cells. Bladder washes yield more cells than void urines. Currently, the company UroCor, Inc. and other firms provide services to diagnose disease. Bladder cancer cells collected in

-70-

urine are examined for morphology, DNA content and for protein
biomarkers as detected immunochemically. While these tests are not
as informative as desired, it is clear that cancer cells can be
collected from patient urines and that these cells are relatively

5   intact with respect to morphology, nucleic acid and protein
content, and expression. RNA has not been directly examined in
these cells, but it is reasonable to expect that the RNA content of
these cells will also be intact relative to the needs of the
envisioned NAP-protection assay.

10   A NAP-protection assay can be designed to determine the
relative abundance of mRNAs for four surrogate biomarkers of
bladder cancer and/or its progression relative to the abundance of
a mRNA encoding a "house keeping" gene that is not differentially
regulated. Changes in the abundance of the surrogate biomarkers

15   relative to the undifferentially expressed gene will be informative
in both diagnosing and evaluating the state of a patient's bladder
cancer. In one embodiment, the final assay uses five probes, each
specific for a different gene. One of the probes will hybridize to
an appropriate "house keeping" gene. The hybridization signal from

20   this probe will act as one of the positive controls in the assay.
Of the remaining four mRNA species, targets will be selected such
that in cancerous cells two of the mRNAs are increased in abundance
and two are decreased relative to nontransformed cells. In this
format, each regulatory pathway acts as an internal control for the

25   other.

-71-

The following assays can be carried out to probe for mRNA levels of the human autocrine motility factor receptor, which is differentially expressed in bladder cancer cells. The DNA sequence for this gene is shown in Figure 7.

5   Ribonuclease NAP Protection Assay

One ribonuclease assay is based upon the currently utilized radioisotopic RNase protection assay, described above. However, the radioactively labeled nucleotide substrates are replaced by 5-S-R-XTP nucleotides, where X is U, dU, C, or dC and R is one of

.0   several different reporter groups. The NAPs for the assay are ssRNA, which are synthesized enzymatically utilizing a bacterial or phage RNA polymerase.

In this assay, RNA NAPs can be prepared in which the RNA is labeled with many fluorescent, colormetric, chemiluminescent, or

l5   antigenic reporter groups. One such antigenic reporter group is the dinitrophenyl group (DNP), and the analog for these experiments is proposed below. The assay is the same as the current method previously described, but the protected RNA products are detected by either monitoring fluorescent emission, probing with antibodies,

?0   or monitoring chemiluminescence. In this assay, final detection still requires the analysis of the NAP protection assay by gel electrophoretic separation of the products. However, the probes can be synthesized in large quantities and stored, so this assay still has a major advantage over the existing methodologies.

-72-

Deoxyribonuclease NAP Protection Assay

An alternative nuclease protection assay can utilize DNA NAPs, similarly modified with multiple reporter groups. One compelling reason to consider DNA NAPs is that they are both chemically and

5   biologically more stable than the comparable RNA NAPs. This makes them better prospects for synthesis in large quantities and for long storage. Despite this inherent stability relative to RNA, ssDNA can be easily degraded by single-strand specific DNases, in the manner as utilized for ssRNA digestion. In addition, with DNA

10   NAPs, one can synthesize large quantities of the DNA probe by utilizing the DNA amplification method, PCR (polymerase chain reaction).

1)   DNA NAPs Prepared with Analog-Tagged Oligonucleotide
      Primers

15   Already developed for use in such a DNA NAP-protection assay is the analog 5-DNP-dU phosphoramidite. This analog can be incorporated chemically into ssDNA oligonucleotides utilizing automated synthesis. Therefore oligonucleotide primers for the PCR reaction can be made which contain multiple analogs, with nearly

20   every "T" in the oligonucleotide synthesis and isolation, and we have shown that it is also retained during the thermal cycling reaction of PCR. Therefore, one can prepare DNA NAPs which contain several reactive thiol groups in the 5' region of the NAP, and these can then be deprotected and modified with a variety of

25   alkylating agents. One can therefore prepare probes for several

-73-

different mRNA targets, each with a different reporter group, and differentially quantify the levels of each in one reaction. This aspect of the method becomes important for the progression of the assay to automation.

5   2)   DNA NAPs Prepared with Analog-Tagged Deoxynucleoside
                 Triphosphates

DNA NAPs can be synthesized by incorporating several (e.g., five to ten) 5-S-DNP-dUs into the 5'-tagged oligonucleotide, and then synthesizing full length DNA NAP by PCR with unmodified dNTPs,
10   e.g., 5-S-DNP-dUTP.

Automation of NAP Protection Assays

Automation of either the RNA NAP Protection Assay or the DNA NAP Protection Assay requires that the electrophoresis step be eliminated. The separation of the individual protected mRNA
15   species by electrophoresis is one of the more time-consuming steps of the current assay, and makes this method intractable to automation. In order to eliminate this step, one must devise a method for the detection and relative quantitation of more than one mRNA species in a single reaction. Currently, the best approach
20   for this involves the use of a variety of fluorescent tags, each with different excitation and emission properties. An increasing number of alkylating reagents tagged with fluorescent groups such as fluorescein or rhodamine, to name only two of many, are now commercially available. Companies are designing fluorescent

-74-

detectors which can differentiate among the different signals produced by these fluorescent dyes. This approach is modeled upon the automated assays currently used for nonisotopic DNA sequencing. The final choice of fluorescent dye combinations used for these

5   assays are determined empirically during the course of the first year of this grant. Alternatively, one can label different NAPs with different antigenic groups, so that RNAs could be distinguished with antibodies, a method already utilized in many automated assays.

10   Deoxyribonuclease NAP Protection Assay

For the DNA NAP assay, one can incorporate commonly used methods for the immobilization of DNA onto resins or microtiter plates. If, for example, a 96-well microtiter plate were coated with streptavidin, then DNA tagged at the 5' end with biotin could

15   be efficiently immobilized. In this way, the DNA NAP and any RNA to which it hybridizes, could be immobilized on a solid matrix, while other components of the hybridization mixture could be washed away.

The following approach can be followed for establishment of an

20   automated DNase NAP Protection Assay.

1)   DNA oligonucleotides can be synthesized which are tagged at the 5' end with one of many groups that would allow immobilization of the DNA to a solid matrix (for example

-75-

biotinylated nucleotides and streptavidin-agarose resin). The DNA NAPs can be multiply-labeled by either:

a) incorporating several (5-10) 5-S-DNP-dUs into the 5'-tagged oligonucleotide, and then synthesizing full length DNA NAP by PCR with unmodified dNTPs.

b) incorporating multiple (50-100) 5-S-DNP-dUs into the body of the DNA product produced by PCR by using 5-S-DNP-dUTP as a substrate.

2) The 5' end-labeled, internally reporter-labeled DNA NAPs can be deprotected and modified with one of the reporter groups chosen in the studies described above.

3) Reaction mixtures can be prepared containing multiple mRNA targets and one or more reporter-tagged NAPs.

4) After hybridization, reactions mixtures can be treated with nucleases specific for single-stranded nucleic acids (such as S1 or Mung Bean Nucleases), to produce the protected mRNA species.

5) Reactions mixtures can be added to microtiter plates coated with Streptavidin (for 5'-biotin labeled DNA) or other materials which are specific for binding the groups commonly used to tag DNA probes. Since the 5' end of the DNA is protected during the NAP-protection assay (see Fig. 1), and NAP which does not

-76-

hybridize are degraded by nuclease, the only material which should bind to the solid matrix is the DNA-mRNA hybrid.

6)   Microtiter plates can be washed to remove unincorporated nucleotides, thereby removing essentially all background signal.

5   7)   Plates can be analyzed with a fluorescent microtiter plate reader to quantify the relative amounts of mRNAs produced for different NAPs.

-77-

REFERENCES

Agrawal, S., ed. <u>Methods in Molecular Biology</u>, ed. Walker, J.M., Vol. 20, Humana Press: Totowa, New Jersey (1993).

Barone, A.D., Tang, J.-Y. and Caruthers, M.H. (1984) *In situ* Activiation of bis-Dialkylaminophosphines - A New Method for Synthesizing Deoxyoligonucleotides on Polymer Supports *Nucleic Acids Res.* 12, 4051-4061.

Bartholomew, B., Braun. F.R., Kassavetis, G.A. and Geiduschek, E.P. (1994) Probing Close DNA Contacts of RNA Polymerase III Transcription Complexes with the Photoactive Nucleoside 4-Thiodeoxythymidine *J. Biol. Chem.* 269, 18090-18095.

Bartholomew, B., Kassavetis. G.A., Braun, B.R. and Geiduschek, E.P. (1990) The Subunit Structure of *Saccharomyces cerevisiae* Transcription Factor IIIC Probed with a Novel Photocrosslinking Reagent *EMBO J.* 9, 2197-2205.

Beaucage, S.L. and Iyer, R.P. (1993) The Functionalization of Oligonucleotides Via Phosphoramidite Derivatives *Tetrahedron* 49, 1925-1963.

Bergstrom, D.E., Beal, P., Jenson, J. and Lin, X. (1991) Palladium-mediated Synthesis of C-5 Pyrimidine Nucleoside Thioethers from Disulfides and Mercurinucleosides *J. Org. Chem.* 56, 5598-5602.

Bleasdale, C., Ellwood, S.B. and Golding, B.T. (1990) 4,4'-Dimethoxytrityl and 4-Monomethoxytrityl Tetrafluoroborate: Convenient Reagents for the Protection of Primary Alcohols Including Sugars *J. Chem. Soc. Perkin Trans 1* 803-804.

-78-

Bradley, D.H. and Hanna, M.M. (1992) Synthesis and Utility of 5-
     Thiocyanato Dexoyuridine and Uridine Phosphoramidites as
     Masked Synthons *Tetrahedron Lett.* 33, 6223-6226.

Brown, H.C., Narasimhan, S. and Choi, Y.M. (1982) Selective
5   Reduction. 30. Effect of Cation and Solvent on the Reactivity
     of Saline Borohydrides for Reduction of Carboxylic Esters.
     Improved Procedures for the Conversion of Esters to Alcohols
     by Metal Borohydrides *J. Org. Chem.* 47, 4702-4708.

Chaudhary, S.K. and Hernandez, O. (1979) A Simplified Procedure for
10   the Preparation of Triphenylmethyleters *Tetrahedron Lett.* 95-
     98.

Christopherson, M.S. and Broom, A.D. (1991) Synthesis of
     Oligonucleotides Containing 2'-Deoxy-6-thioguanosine at a
     Predetermined Site *Nucleic Acids Res.* 19, 5719-5724.

15 Clegg, R.M., Murchie, A.I.H., Zechel, A., Carlberg, C., Diekmann,
     S. and Lilley, D.M.J. (1992) Fluorescence Resonance Energy
     Transfer Analysis of the Structure of the Four-Way DNA
     Junction *Biochemistry* 31, 4846-4856.

Clivio, P., Fourrey, J.-L., Gasche, J., Audic, A., Favre, A.,
20   Perrin, C.; and Woisard, A. (1992) Synthesis and Purification
     of Oligonucleotides Containing Sulfur Substituted Nucleobase:
     4-Thiouracil, 4-Thiothymine and 6-Mercaptopurine *Tetrahedron
     Lett.* 33, 65-68.

Coleman, R.S. and Kesicki, E.A. (1994) Synthesis and Postsynthetic
25   Modification of Oligodeoxynucleotides Containing 4-Thio-2'-
     deoxyuridine ($d^{S4}U$) *J. Am. Chem. Soc.* 116, 11636-11642.

-79-

Connolly, B.A. and Newman, P.C. (1989) Synthesis and Properties of Oligonucleotides Containing 4-Thiothymine, 5-Methy-2-pyrimidinone-1-β-D-(2'-deoxyriboside) and 2-Thiothymidine *Nucleic Acids Res. 17*, 4957-4974.

5   Cosstick, R. and Douglas, M.E. (1991) Synthesis of a Dinucleoside Monophosphate Analogue Containing 6-N-(2-Aminoethyl)-2'-Deoxyadenosine. A Novel Approach to Sequence Specific Crosslinking in Synthetic Oligonucleotides *J. Chem. Soc. Perkin Trans. 1* 1035-1040.

10   Dissinger, S. and Hanna, M.M., (1990) J. Biol. Chem., 265:7662.

Flaschka, H.A., Barnard, A.J. and Sturrock, P.E. (1969) *Quantitative Analytical Chemistry, Short Introduction to Practice*, Barnes & Noble, INC., New York.

Gait, M.J., ed. Practical Approach Series, ed. Rickwood, D. and

15   Harnes, B.D., IRL Press: Oxford (1984).

Gao, H., Fathi, R., Gaffney, B.L., Goswami, B., Kung, P.-P., Rhee, Y., Jin, R. and Jones, R.A. (1992) 6-O-(Pentatluorophenyl)-2'-deoxyguanosine: A Versatile Synthon for Nucleoside and Oligonucleotide Synthesis *J. Org. Chem. 57*, 6954-6959.

20   Glick, G.D. (1991) Synthesis of a Conformationally Restricted DNA Hairpin *J. Org. Chem. 56*, 6746-6747.

Gibson, K.J. and Benkovic, S.J. (1987) Nucleic Acid Res., 15, No. 16:6455.

Goodwin, J.T. and Glick, G.D. (1993) Incorporation of Alkylthiol

25   Chains at C-5 of Deoxyuridine *Tetrahedron Lett. 34*, 5549-5552.

Goodwin, J.T., Osborne, S.E., Swanson, P.C. and Glick, G.D. (1994) Synthesis of a Disulfide Cross-linked DNA Triple Helix *Tetrahedron Lett.* 35, 4527-4530.

Greene, T.W. (1981) *Protective Groups in Organic Synthesis*, John Wiley & Sons, New York.

Grzybowski, J., Will, D.W., Randall, R.E., Smith, C.A. and Brown, T. (1993) Synthesis and Antibody-mediated Detection of Oligonucleotides Containing Multiple 2,4-Dinitrophenyl Reporter Groups *Nucleic Acids Res.* 21, 1705-1712.

Hanna, M.M. (1989) Methods in Enzymology, 180:383

Hanna, M.M. (1993) Photocrosslinking Analysis of Protein-RNA Interactions in *E. coli* Transcription Complexes *Cellular Molecular Biology Research* 39, 393-399.

Hanna, M.M., Dissinger, S., Williams, B.D. and Colston, J.E. (1989) Synthesis and Characterization of 5-[4-(Azidophenacyl)thio]uridine 5'-Triphosphate, a Cleavable Photo-Cross-Linking Nucleotide Analogue *Biochemistry* 28, 5814-5820.

He, B. and Hanna, M.M., (1995) Nucleic Acids Res., 23.

Hershey, A.D. and Krause, J.E., (1989) Science, 247:960.

Jadhav, V., Barawkar, D., Natu, A., and Ganesh, K. *Nucleosides and Nucleotides*, (1997) 16, 107-114.

Kalman, T.I. and Bardos, T.J. (1967) The Autoxidation of 5-Mercaptouracil and 5-Mercaptodeoxyuridine *J. Am. Chem. Soc.* 89, 1171-1175.

-81-

Keller, G.H., and Manak, M.M., (1989) <u>DNA Probes</u>, New York, N.Y.: Stockton Press.

5   Kini, G.D., Robins, R.K. and Avery, T.L. (1989) Synthesis and Antitumor Activity of Ribavirin Imidates.   A New Facile Synthesis of Ribavirin Amidine (1-β-D-Ribofuranosyl-1,2,4-triazone-3-carboxamidine Hydrochloride) *J. Med. Chem.* 32, 1447-1449.

Kricka, L.J. (1992) ed. <u>Nonisotopic DNA Probe Techniques</u>, Academic Press, Inc.: San Diego, CA.

10   Kuimelis, R.G. and Nambiar, K.P. (1994) Synthesis of Oligodeoxynucleotides Containing 2-Thiopyrimidine Residues- A New Protection Scheme *Nucleic Acids Res.* 22, 1429-1436.

Lakshman, M.K., Sayer, J.M., Yagi, H. and Jerina, D.M. (1992) Synthesis and Duplex-Forming Properties of a Nonanucleotide
15   Containing an N^6-Deoxyadenosine Adduct of a Bay-Region Diol Epoxide *J. Org. Chem.* 57, 4585-4590.

Maurizi, M.R. and Ginsburg, A. (1986) <u>Biochemistry</u>, 25:131.

Mergny, J.-L., Boutorine, A.A., Garestier, T., Belloc, F., R ougee, M., Bulychev, N.V., Koshkin, A.A., Bourson, J., Lebedev, A.V.,
20   Valeur, B., Thuong, N.T. and Helene, C. (1994) Fluorescence energy transfer as a probe for nucleic acid structures and sequences *Nucleic Acids Res.* 22, 920,928.

Nagamachi, T., Fourrey, J.-L., Torrence, P.F., Waters, J.A. and Witkop, B. (1974) Synthesis, Chemistry, and Biological
25   Activity of 5-Thiocyanatopyrimidine Nucleosides as Potential Masked Thiols *J. Med. Chem.* 17, 403-406.

-82-

Scremin, C.L., Zhou, L., Srinivasachar, K. and Beaucage, S.L. (1994) Stepwise Regeneration and Recovery of Deoxyribonucleoside Phosphoramidite Monomers during Solid-Phase Oligonucleotide Synthesis *J. Org. Chem.* *59*, 1963-1966.

5   Shaltiel, S. (1967) Thiolysis of Some Dinitrophenyl Derivatives of Amino Acids *Biochem. Biophys. Res. Comm.* *29*, 178-183.

Sheardy, R.D. and Seeman, N.C. (1986) The Synthesis of a Deoxyoligonucleotide Incorporating 5-Iododeoxyruidine *J. Org. Chem.* *51*, 4301-4303.

10   Sinha, N.D., Biernat, J., McManus, J., and Koster, H. (1984) Polymer Support Oligonucleotide Synthesis XVIII: Use of β-cyanoethyl-*N,N*-dialkylamino-/*N*-morpholino Phosphoramidite of Deoxynucleosides for the Synthesis of DNA Fragments Simplifying Deprotection and Isolation of the Final Product *Nucleic Acids Res.* *12*, 4539-4557.

15

Soai, K. and Ookawa, A. (1986) Mixed Solvents Containing Methanol as Useful Reaction Media for Unique Chemoselective Reductions with Lithium Borohydride *J. Org. Chem.* *51*, 4000-4005.

Torrence, P.F., Waters, J.A. and Witkop, B. (1968) 2'Deoxy-5'(thiocyanato)uridine [1-(2-deoxy-β-D-erythropentofuranosyl)-5-(thiocyanato)uracil] *Synthetic Procedures in Nucleic Acid Chemistry*, (W.W. Zorbach and R.S. Tipson, Ed.) pp 367-370, Interscience, New York.

Trivedi, B.K. and Bruns, R.F. (1989) C2.N^6-Disubstituted Adenosines: Synthesis and Structure-Activity Relationships *J. Med. Chem.* *32*, 1667-1673.

-83-

Waters, T.R. and Connolly B.A. (1992) Straightforward Synthesis of 6-Thiodeoxyguanosine and Its Incorporation into Oligodeoxynucleotides *Nucleosides Nucleotides* 11, 985-998.

Xu, Y.-Z., Zheng, Q. and Swann, P.F. (1992a) Synthesis and Duplex Stability of Oligodeoxynucleotides Containing 6-Mercaptopurine *Tetrahedron Lett.* 33, 5837-5840.

Xu, Y.-Z., Zheng, Q. and Swann, P.F. (1992b) Synthesis by Post-Synthetic Substitution of Oligomers Containing Guanine Modified at the 6-Position with S-, N-, O-derivatives *Tetrahedron* 48, 1729-1740.

Xu, Y.-Z., Zheng, Q. and Swann, P.F. (1992c) Synthesis of DNA Containing Modified Bases by Postsynthetic Substitution. Synthesis of Oligomers Containing 4-Substituted Thymine: O^4-Alkylthymine, 5-Methylcytosine, N^4-(Dimethylamino)-5-methylcytosine, and 4-Thiothymine *J. Org. Chem.* 57, 3839-3845.

Yang, X.J., Hart, C.M., Grayhack, E.J. and Roberts, J.W. (1987) Transcription Antitermination by Phage Lambda Gene Q Protein Requires a DNA Segment Spanning the RNA Start Site *Genes Dev* 1, 217-226.

Zhang, Y. and Hanna, M.M. (1995) Expression and Functional Characterization of *Escherichia coli* NusA and Lambda Q as Glutathione-S-Transferase Fusion Proteins *Protein Expression and Purification* 6, 625-631.

Changes may be made in the compounds described herein or in
the steps or the sequence of steps of the methods of synthesis or
use of the compounds described herein without departing from the
spirit and scope of the invention as defined in the following
5   claims.

What is claimed is:

1. A nucleotide analog of the formula:

wherein

R_1 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R_2 is -H, -OH, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, -OR$_4$, or a solid support bonded via an O at the 3' position;

R_3 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R_4 is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at the 5 position on said base that is not involved in Watson-Crick base pairing or does not disrupt normal Watson-Crick base pairing, said protected thiol group being stable under conditions of chemical nucleic acid synthesis and/or conditions of enyzmatic nucleic synthesis and being convertible to a reactive thiol after said synthesis.

-86-

2.   The nucleotide analog of Claim 1 wherein said base is selected from the group consisting of cytosine and uracil.

3.   The nucleotide analog of Claim 1 wherein said base compromises an additional protecting group on a reactive moiety of said base.

4.   The nucleotide analog of Claim 3 wherein said base is selected from the group consisting of N^4-anisoyl cytosine, N^4-benzoyl cytosine, N^4-isobutyryl cytosine and N^4-acetyl cytosine.

5.   The nucleotide analog of Claim 1 wherein said protected thiol group on said base is selected from the group consisting of thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R_{10}-DNP), and alkyldisulfide (-S-S-R_{10}) wherein R_{10} is a lower alkyl.

6.   The nucleotide analog of Claim 1, wherein the protecting group of R_1 is selected from the group consisting of p-(dimethoxytrityl), p-(monomethoxytrityl), fluorenylmethyl-oxycarbonyl, levuloyl and 9-phenylxanthene-9-yl.

7.   The nucleotide analog of Claim 1, wherein protecting group of R_3 is selected from the group consisting of 1-(2-chloro-4-methylphenyl)-4-methoxy-4-piperidinium, 2'-acetal, o-nitrobenzyl, tert-butyldimethyl silyl, tetrahydrofuranyl and 4-methoxytetrahydropyranyl.

5

-87-

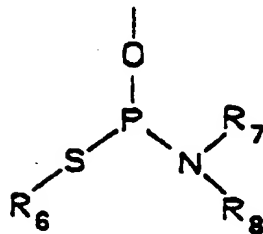8. The nucleotide analog of Claim 1 wherein R_1 is $-OR_4$ wherein R_4 is a protecting group and R_2 is a phosphoramidite group.
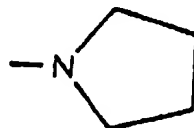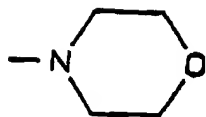
9. The nucleotide analog of Claim 8 wherein said base is cytosine, or uracil.

10. The nucleotide analog of Claim 9 wherein said base comprises an additional protecting group on a reactive moiety of said base.

11. The nucleotide analog of Claim 12 wherein said base is selected from the group consisting of N^4-anisoyl cytosine, N^4-benzoyl cytosine, N^4-isobutyryl cytosine and N^4-acetyl cytosine.

12. The nucleotide analog of Claim 8 wherein said protected thiol group of said base is selected from the group consisting of thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R_{10}-DNP), and alkyldisulfide (-S-S-R_{10}) wherein R_{10} is a lower alkyl.

13. The nucleotide analog of Claim 8 wherein the protecting group of R_1 is selected from the group consisting of p-(dimethoxytrityl), p-(monomethoxytrityl), fluorenylmethyl-oxycarbonyl, levuloyl, and 9-phenylxanthene-9-yl.

14. The nucleotide analog of Claim 8, wherein the protecting group of R_3 is selected from the group consisting of 1-(2-chloro-4-

-88-

methylphenyl)-4-methoxy-4-piperidinyl, 2'-acetal, o-nitrobenzyl,
tert-butyldimethyl silyl, tetrahydrofuranyl and 4-
methoxytetrahydropyranyl.

15. The nucleotide analog of Claim 1 or 8, wherein said phosphoramidite of R_2 is represented by the formula:

$$
\begin{array}{c}
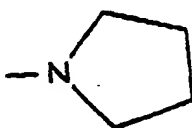| \\
O \\
| \\
P \\
\diagup \quad \diagdown \\
O \qquad N \diagup R_7 \\
| \qquad\qquad \diagdown R_8 \\
R_6
\end{array}
$$

R_6 is a lower alkyl, cyanoethyl or a substituted lower alkyl; and

R_7 and R_8 are independently lower alkyls, or when taken together with the nitrogen to which they are attached comprise one of the groups:

16. The nucleotide analog of Claim 1 wherein R_1 is $-OR_4$ wherein R_4 is a protecting group and R_2 is a phosphorothioamidite group.

-89-

17. The nucleotide analog of Claim 1 or 16, wherein said R_2 is a phosphorothioamidite represented by the formula:
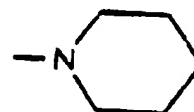

wherein

R_6 is a lower alkyl, cyanoethyl or a substituted lower alkyl; and

R_7 and R_8 are independently lower alkyls, or when taken together with the nitrogen to which they are attached comprise one of the groups:

18. The nucleotide analog of Claim 1 wherein R_2 is an O-substituted monophosphate group selected from the group consisting of -2-chlorophenyl monophosphate, O-2,5-dichlorophenyl monophosphate, O-2,2,2-trichloroethyl monophosphate and the N oxide of 4-methoxypyridine-2-methylene monophosphate.

-90-

19.   A nucleotide analog of the formula:

wherein

R_1 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

5   R_2 is a phosphoramidite group;

R_3 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R_4 is a lower alkyl or a protecting group; and

B   is a modified pyrimidine base comprising a

10   thiodinitrophenyl group attached at the 5 position

on said pyrimidine base.

20.   The nucleotide analog of Claim 19 wherein said base is selected from the group consisting of N^4-anisoyl cytosine, N^4-benzoyl cytosine, N^4-isobutyryl cytosine and N^4-acetyl cytosine.

21.   The nucleotide analog of Claim 19, wherein R_1 is -OR$_4$ and R_4 is a protecting group selected from the group consisting of p-(dimethoxytrityl), p-(monomethoxytrityl), fluorenylmethyl-oxycarbonyl, levuloyl, and 9-phenylxanthene-9-yl.

22.   The nucleotide analog of Claim 19 wherein said phosphoramidite of R_2 is represented by the formula:

$$O - P \begin{matrix} O - R_6 \\ N \begin{matrix} R_7 \\ R_8 \end{matrix} \end{matrix}$$

wherein

R_6 is a lower alkyl, cyanoethyl or substituted lower alkyl; and

R_7 and R_8 are independently lower alkyls, or when taken together with the nitrogen to which they are attached form the groups:
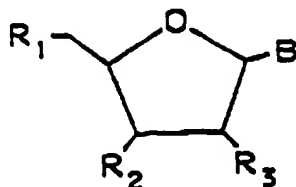
$$-N\bigcirc O \quad , \quad -N\bigcirc \quad , \text{ or } \quad -N\bigcirc .$$

23.   The nucleotide analog of Claim 22 wherein R_6 is 2-cyanoethyl, R_7 is isopropyl and R_8 is isopropyl.

24.   The nucleotide analog of Claim 23 wherein said base is uridine or deoxyuridine.

25.   The nucleotide analog of Claim 24 wherein R_1 is $-OR_4$ and R_4 is p-(dimethoxytrityl).

26.   A nucleotide analog of the formula:

wherein:

R$_1$ is -H, -OH, a mono, di, or triphosphate group, or
       -OR$_4$;

R$_2$ is -H, -OH, a mono, di, or triphosphate group, a
       phosphoramidite group, a phosphorothioamidite
       group, a phosphonate group, an O-substituted
       monophosphate group, -OR$_4$, or a solid support bonded
       via an O at the 3' position;

R$_3$ is -H, -OH, a mono, di, or triphosphate group, or
       -OR$_4$;

R$_4$ is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected
       thiol group attached at the 5 position on said
       base.

27.   The nucleotide analog of Claim 26 wherein said base is
cytosine, or uracil.

28.   The nucleotide analog of Claim 26 wherein said protected
thiol group on said base is selected from the group consisting of
thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R$_{10}$-DNP), S-
phenylacetamidomethyl(-S-CH$_2$NHCOCH$_2$Ph) wherein R$_{10}$ is a lower alkyl.

-93-

29. The nucleotide analog of Claim 26, wherein R₁ is a monophosphate, a diphosphate, or a triphosphate.

30. An oligonucleotide containing the nucleotide analog of any one of Claims 1, 8, 16, 19, or 26.

31. A method of producing the nucleotide analog of any one of Claims 1, 8, 16, 19, or 26 which comprises preparing a thiol-protected nucleoside or nucleotide base wherein said thiol is attached to the 5 position on said base; reacting said nucleoside or nucleotide base under conditions to effect conversion of said base to a phosphoramidite, phosphorothioamidite, phosphonate, o-substituted monophosphate or phosphate nucleotide analog and under conditions which do not destroy the protected thiol; and recovering said analog.

32. A method of synthesizing a nucleic acid having an attached functional group which comprises incorporating a thiol-protected nucleotide analog of any one of Claims 1, 8, 16, 19, or 26 into a nucleic acid by a chemical or enzymatic method of nucleic acid synthesis; recovering said nucleic acid containing said analog; deprotecting the analog of said nucleic acid to produce a nucleic acid containing a reactive thiol group; treating the reactive thiol group with a thiol modifying reagent to thereby attach a functional group and produce said nucleic acid with an

-94-

10   attached functional group; and recovering said nucleic acid with
said attached functional group.

33.   The method of Claim 32 wherein said functional group is
selected from the group consisting of a photocrosslinker, a
crosslinker, a reporter molecule, a radioisotope, a fluorescent
group, a spin label, chemiluminescent or an antigenic group.

34.   The method of Claim 33 wherein said photocrosslinker is
an aryl azide.

35.   The method of Claim 33 wherein said reporter group is
biotin, an enzyme, or a fluorescent molecule.

36.   The method of Claim 33 wherein said fluorescent group is
fluorescein.

37. A method for producing a deoxyuridine analog having a
protected thiol group on the 5 position, comprising:
         reacting a diacetyl 5-thiocyanate deoxyuridine to form a
            5-dinitrophenylthio deoxyuridine;

5   converting said 5-dinitrophenylthio deoxyuridine to a
            phosphoramidite, phosphorothioamidite, phosphonate,
            O-substituted monophosphate or phosphate nucleotide
            analog under conditions which do not destroy the 5-
            dinitrophenylthio group;

-95-

10

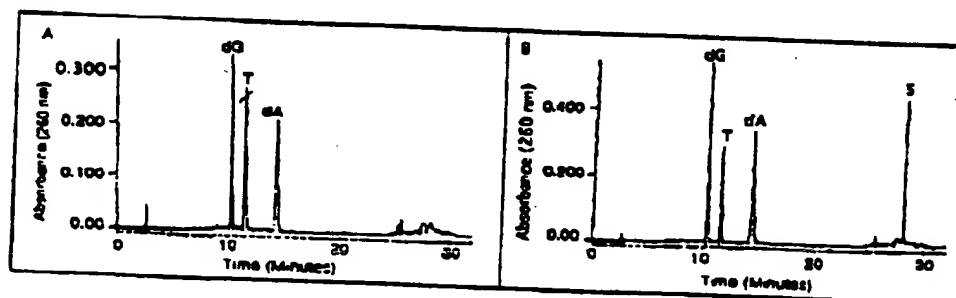and recovering said 5-dinitrophenylthio deoxyuridine
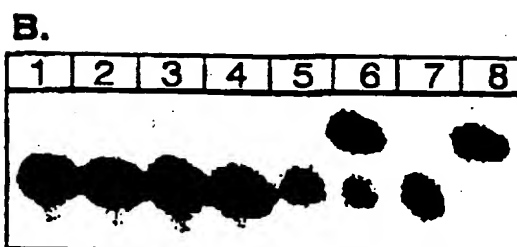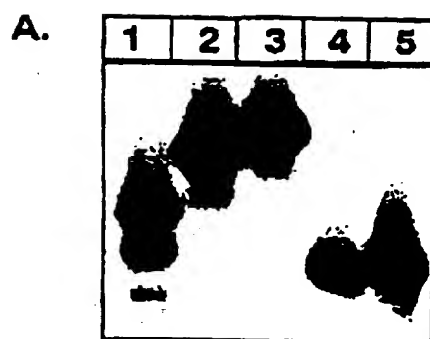analog.

A.

B.

S-DNP
|
5' OH-dU(pN)$_n$-OH 3'

1. Radioactive end-labeling of oligonucleotide →

S-DNP
|
5' ^{32}P-dU (pN)$_n$-OH 3'

2. Primer annealing to ss DNA

^{32}P
S-DNP
OH

3. Primer extension of analog-tagged oligonucleotide ←

^{32}P
S-DNP

4. Deprotection and Addition of crosslinker

^{32}P
S-Ar-N$_3$

5. DNA-Protein Crosslinking →

^{32}P
S-Ar-

6. Nuclease treatment

^{32}PO$_3^-$
O
drib-dU-S-Ar-

Figure 6: Ribonuclease Protection Assay

Hatched areas indicate regions of RNA-RNA homology with the target gene (hybridization and double-stranded structure). Shaded areas in the RNA probe are not homologous to the target mRNA and therefore do not hybridize, leaving regions of single-stranded RNA. These ssRNA regions are susceptible to cleavage by certain nucleases.

FIGURE 7: DNA Sequence of Human Autocrine Motility Factor Receptor (AMFR)

```
   1 GAAACTGCCT GTGAACATCT TTTCCACAACT CCTGTCTTC GTTCCTGGCT AGAACAAGAC
  61 ACCTCCTGTC CAACATGCAG AATGTCTCTT AATATTGCCG ACAATATCGT GTTCAGGGAAG
 121 AACATCAAGG AGAGAACTTG GATGAGAATT TGGTTCCTGA TGCAGCAGCC GAAGGAGAACC
 181 TCGCTTAAAC CAACACAATC ACTTCTTCCA TTTGGATGGG TCTCGGATTG CGAGCTGGCT
 241 GCCGAGTTTT TCGGTTGAAG TGATGCACAC CACCAACATT CTTGGCATTA CGCAGGCCAG
 301 CAACTCCCAG CTCAATGCAA TGGCTCATCA GATTCAAGAG ATGTTTCCCC AGGTTCCATA
 361 CCATCTGGTA CTTCCAGGACCC TCCAGCTGAC AGGCTCAGTT AATAACAAC AGTCAATATT
 421 TTAGAAGGAC GGATTCAACT ACCTTTTCCT ACACAGCGGT CAGATAGCAT CAGAACTGCA
 481 TTGAACAGTC CTGTGTGAAAG GCCAAGCCAGT GACCAGGAAG AGGGAGAAAAC TTCTGTCTCAG
 541 ACCGAGCGTG GTGCCACTGG ACCTCAGTCC TCGCCTGGAG GAGAGCGTCG ACTTCGGCGGA
 601 GGTGGAAGTG GAGCCCAGTG AGGTGGAAGA CTTCGAGCGT CCGTTGGGAGCC GCTTCTCCAA
 661 GTCTGCTGAT GAGAGAGACAGC GCATGCTCGT CGCCAGCGGTAA GGACGAACTC CTCCAGCAAG
 721 CTCCCAAACG TTTCTTGAAC AAAAGTTCTG AAGATGATCC GCCCTCAGAG AGCTTCCTCC
 781 CCTCGGAAGG TGCGTCCTCT GACCCCGGTGA CCCGTCGGTCG AAGGATGCTG GCTGCGCGGGC
 841 GGAACGGAGG CTTCAGAAGCC GAGCAGAGACCTC CTAGCCGCTCC CTTGCCTTCC TCAGCTGCCT
 901 CCTCGCCGCCT GTGCCCGACT GACTCGGAGGA GGCCTGTCCC AATTCTGCCG CTCCATGGAA
 961 AGCCCGGGCTT CACTGCCATTG CCGCTCTGTATA AAGCATGTGG TCTTATAGTG TTTGGACAGC
1021 TGATAAATTT AATCCTTCTT TGTAATACTT TCTATGTGAC ATTTCTCTTCCCCTTAGAAA
1081 CACTGCAAAT TTTAACTGTA GGTATGATCT CTTCTGGTGT TGACTGGACT GCTTGGCGGTG
1141 GGGGAGGATC AGGAGGAAGT GAGCCCAGTCG CCTGCCTGCA GCAGGCAGCT TCTACTCCTG
1201 CCTCCATGCAT AGGTCCCACA AATGCAGGTG TCCTGAGCAC CACACCCAGT GGGAAGAGTG
1261 TGGGGGAGGC GCACAGTGTG AGCCCGCGCCC CACGTCGTGG GGTAACATCT GTTATCAAAC
1321 TGTCTGTGTT GTTGTGGAAG CATGTAGACT GTGCCAGAGC CAGACCCACG GGCTCATGCA
1381 CCCTGAGCA GCAGCGGCATC TTGGAAAAGG AACTCTTGGT TCGATACCTG GAGCAGAGGA
1441 GGGGAAAGTC CAGGGCTATA GGGTTGTGATG AAGTCACCCC TTTCTGTGTCC ACTACACTGTG
1501 GGACTGACTT TCCGAGCCTC CAGTCCAAAG CCGGCTTGAT TTCCGTGAAC TCTGGTGCTC
1561 CTCCATCTCA TCAGTGTGCCC CCATGGGTCC CCTCCCTCT CAGCCATTTCC TTGTCCCGTC
1621 TCGGACCTGGG GAGTGGTTAG GCAAGCAAGCT TTGGCTTATG GTTTCATTC ATTGGTGAAG
1681 TAAATTAGGC TAGTTGCACGT AAAGCCTGTG GGTTTTGGTCC TTGAACAAGA TGTGTGGGCTT
1741 GCAAGATGCGG AGAGTAAACC TTAAG
```

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6   C07H21/00   C07H19/10   C07H19/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6   C07H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category°	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 01445 A (RES CORP TECHNOLOGIES INC) 20 January 1994 see the whole document ---	1-4, 6-11, 13-18, 26,27, 29-36
X	M.M. HANNA: "Photochemical Cross-Linking Analysis of Protein-Nucleic Acid Interactions in Escherichia coli Transcription Complex from Lambda PR' Promoter" METHODS IN ENZYMOLOGY, vol. 274, 1996, pages 403-418, XP002045925 see the whole document --- -/--	1-37

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 November 1997	1 6. 12. 97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bardili, W

Form PCT/ISA/210 (second sheet) (July 1992)

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	M. M. HANNA AND K.L. MEYER: "Synthesis and Characterization of a New 5-Thiol-Protected Deoxyuridine Phosphoramidite for Site-Specific Modification of DNA" BIOCONJUGATE CHEMISTRY, vol. 7, no. 4, 1996, pages 401-412, XP002045926 see the whole document ---	1-37
A	J. CHATTOPADHYAYA ET AL.: "Synthesis of C-5 and N-3 Arenesulfenyl Uridines" ACTA CHEM. SCAND. SER. B, vol. 39, 1985, pages 203-12, XP002045927 -----	

1

BNSDOCID: <WO 9803532A1 I >

INTERNATIONAL SEARCH REPORT

Information on patent family members

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9401445 A	20-01-94	NONE	

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: BASE-PROTECTED NUCLEOTIDE ANALOGS WITH PROTECTED THIOL GROUPS

(57) Abstract

The present invention is directed to protected thiol analogs of pyrimidine bases for syntheses of DNA and RNA by chemical or enzymatic methods. The subject analogs include reagents suitable for DNA or RNA synthesis via phosphoramidite, H-phosphonate or phosphotriester chemistry as well as reagents suitable for use by RNA and DNA polymerase, including thermostable polymerases employed by PCR or other nucleic acid amplification techniques. The nucleotide analogs synthesized by methods of this invention can thus be incorporated into oligonucleotides or polynucleotides, deprotected and derivatized with a functional group. In some cases the protecting groups are themselves antigenic and may be left on the oligonucleotides or polynucleotides for detection with antibodies. A method of synthesizing oligonucleotides with a functional group using the subject nucleotide analogs is also provided.

BASE-PROTECTED NUCLEOTIDE ANALOGS
WITH PROTECTED THIOL GROUPS

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT
Not Applicable

5   BACKGROUND OF THE INVENTION
        The present application claims the benefit of the filing
date of U.S. Provisional Application 60/022,573, filed July 24,
1996.
        The present invention is directed to thiol-protected
10   pyrimidine nucleotide analogs which can be used, as one example,
for syntheses of DNA and RNA by chemical or enzymatic methods.
The subject analogs include reagents suitable for DNA or RNA
synthesis via phosphoramidite, H-phosphonate or phosphotriester
chemistry as well as reagents suitable for use by RNA and DNA
15   polymerases, including thermostable polymerases employed by PCR
or other nucleic acid amplification techniques. Methods of
synthesizing the nucleotide analogs are also provided by the
present invention. The nucleotide analogs of this invention can
thus be incorporated into oligonucleotides or polynucleotides,
20   deprotected, and then derivatized with a functional group.
        Oligonucleotides with a variety of modifications have
widespread utility for many purposes, such as stabilizing
oligonucleotides to degradation, introducing reporter groups,
allowing site-specific delivery of therapeutics, and introducing
25   crosslinkers.   Such modifications can occur as modified
internucleotide phosphate linkages or analogs of such linkages,
modified sugars or modified bases. Additionally, 5'- or 3'-end
conjugates of the oligonucleotides represent another class of
modified oligonucleotides. The present invention relates to
30   base-modified nucleotide analogs with protected thiol groups;
these analogs are intermediates for chemical or enzymatic
synthesis of oligonucleotides and polynucleotides.
        The synthesis of oligoribonucleotides and oligodeoxyribo-
nucleotides (i.e. oligonucleotides) containing base-modified
35   nucleotides at specific positions provides a powerful tool in
the analysis of protein-nucleic acid or nucleic acid-nucleic
acid interactions.   These oligonucleotides have many other

SUBSTITUTE SHEET (RULE 26)

potential uses, such as the site-directed delivery of therapeutics, utility as anti-sense therapeutics, and utility as diagnostic probes. Nucleotide analogs can be introduced into nucleic acids either enzymatically, utilizing DNA and RNA polymerases, or chemically, utilizing manual or automated synthesis. Preparation of such oligonucleotides by automated synthesis utilizing, for example, phosphoramidite nucleotides allows for incorporation of a broad range of nucleotide analogs without the restraints for specific substrate conformation of the nucleotides that is imposed by most polymerases. Often, nucleotide analogs containing photoreactive crosslinking groups are introduced into oligonucleotides to probe protein-nucleic acid interactions via photocrosslinking (for a partial review, see Hanna, 1989, Methods Enzymol. 180:383-409; and Hanna, 1996, Methods Enzymol, 273 Chapter 31). Deoxyoligonucleo-tides containing 4-thiothymidine have been prepared and used for photochemical crosslinking of proteins directly to the nucleotide bases through the group (Nikiforov et al., 1992, Nucleic Acids Res. 20:1209-14). Similarly, oligonucleotides containing 3-(amino-propyl)-2′-deoxyuridine have been prepared and the amino group subsequently modified with fluorescent, photoactive or other reporter groups (Gibson et al., 1987, Nucleic Acids Res. 15:6455-66). However, the former thiodeoxynucleotide suffers the disadvantage that if it is modified with a thiol modifying reagent the normal Watson-Crick base of the nucleotide is drastically affected, making this and similar analogs generally unsuitable for use in enzymatic nucleic acid synthesis.

Other analogs involving modifications at the C5 position of deoxyuridine have also been previously reported. The thiol-containing analog, 5-thiocyanatodeoxyuridine phosphoramidite, provided a 5-mercaptodeoxyuridine moiety within the oligonucleotide following reduction of the thiocyanate (Bradley & Hanna, 1992). However, the thiocyanato moiety displayed variable stability during synthesis of both the nucleotide and the oligonucleotide and therefore this compound did not represent an ideal analog for incorporation of 5-thiol modified nucleotides into nucleic acids. The syntheses of a series of

3

phosphoramidites containing alkylthiol tethers at the C5 position of deoxyuridine has been reported (Goodwin & Glick, 1993). The thiol groups in these analogs are attached to the ring by either a three, four, or five carbon chain. The presence of the carbon chains makes the minimal distance between the molecular probes and the oligonucleotide greater than that which can be achieved with our analog. In addition, these compounds represent alkyl thiol analogs which have a lower reactivity for modification of the thiol group than the 5-mercaptopyrimidine analog. This is due to an increase in acidity of 5-mercaptopyrimidines (pKa~5-5.6) over alkylthiol moieties (pKa~8-10). Phosphoramidites containing alkylthiol tethers at the N3 position of thymidine have also been prepared, but the position of this modification results in a disruption of the Watson-Crick base pairing. These analogs have been used mainly for preparing disulfide cross-links in DNA for studies involving stem loop and triple helical structures (Glick, 1991; Goodwin et al., 1994).

Several 5-modified deoxyuridine phosphoramidites are commercially available which contain functional groups (i.e. carboxylic acids or alkyl amines) for post-modification following incorporation into the oligonucleotide. Likewise, alkyl thiol ethers of 5-mercaptodeoxyuridine containing protected carboxylic acids and alkyl amines have also been described (Bergstrom et al., 1991). The protected functional groups in these analogs are not attached directly to the ring but are positioned at the end of carbon chains. These groups are not as easily modified as mercaptans: the functional groups formed during post-synthetic modification are limited, and an easily cleavable group is not available. Thiol-containing phosphoramidites for incorporating 4-thiothymidine (Clivio et al., 1992; Xu et al., 1992c), 4-thiodeoxyuridine (Clivio et al., 1992; Coleman & Kesicki, 1994), 2-thiothymidine (Connolly & Newman, 1989; Kuimelis & Nambiar, 1994), 6-thiodeoxyguanosine (Christopherson & Broom, 1991; Waters & Connolly, 1992; Xu et al., 1992b) and 6-thioinosine (Clivio et al., 1992a) have been reported. These analogs can occupy internal positions within an oligonucleotide and can serve as photochemical crosslinkers.

SUBSTITUTE SHEET (RULE 26)

4

However, they cannot be further modified without disrupting Watson-Crick base pairing, and therefore, as photocrosslinking probes, these analogs are only useful for evaluating interactions which occur directly with the nucleotide base.

5 Modification with other molecular probes (e.g., fluorescent tags) would also disrupt Watson-Crick base pairing. In addition, the deprotection of oligonucleotides containing these thiol-modified nucleotides must be carefully monitored to prevent conversion of these analogs to the corresponding oxygen
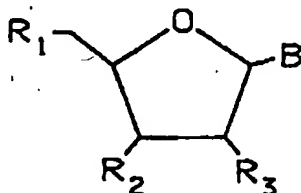
10 and nitrogen derivatives.

Described herein are nucleotide analogs which can be used for site-specific modification of DNA or RNA at internal and terminal positions within the DNA or RNA sequence, and after modification, for molecular probes which can be placed at

15 variable distances from the DNA or RNA backbone.

The present invention provides novel base-protected nucleotide analogs, both ribonucleotides and deoxynucleotides, that contain masked thiol groups on the 5 position of pyrimidines, which is not involved in Watson-Crick base pairing.

20 These analogs can be incorporated into oligonucleotides via automated synthesis and isolated with the thiol protecting group intact. After removal of the thiol protecting group many types of functional groups, such as photocrosslinking agents, fluorescent tags, radioisotopes, biotin, reporter molecules,

25 spin labels (e.g., commercially available proxyl or tempo), chemiluminescent, antigenic or other functional groups, can be site-specifically attached by utilizing thiol-modifying reagents. This feature adds a level of specificity to the oligonucleotide modifications not present with the amino-tagged

30 analogs previously described (Gibson et al, 1987), and enables examination of molecular interactions that are not directly at the nucleotide base by allowing functional groups to be placed at varying distances from the base or helix strand. Since these analogs have the functional group attached via the sulfur atom,

35 some have the further advantage of being cleavable under conditions which will not degrade or modify the oligonucleotide.

SUMMARY OF THE INVENTION

— 5 —

        This invention relates to pyrimidine nucleotide analogs
which contain modified bases with protected thiol groups
attached at a position on the base, preferably the 5 position,
which is not involved in Watson-Crick base pairing.   These
5   nucleotide analogs are intermediates in chemical or enzymatic
synthesis of DNA or RNA oligonucleotides and are therefore
stable under conditions required for synthesis of these
molecules. After synthesis, the protecting group on the analog
is removable to generate a reactive thiol group.   Once
10   generated, the thiol group can be treated with thiol modifying
reagents to attach functional groups such as crosslinking agents
or reporter molecules.

        In particular, the nucleotide analogs of the present
invention have the formula:

15   wherein

        R_1 is -H, -OH, a mono, di, or triphosphate group, or
            -OR_4;

        R_2 is -H, -OH, a mono, di, or triphosphate group, a
            phosphoramidite group, a phosphorothioamidite
20   group, a phosphonate group, an O-substituted
            monophosphate group, -OR_4, or a solid support
            bonded via an O at the 3' position;

        R_3 is -H, -OH, a mono, di, or triphosphate group, or
            -OR_4;

25   R_4 is a lower alkyl or a protecting group; and

        B is a modified pyrimidine base comprising a protected
            thiol group attached at the 5 position on said
            base that is not involved in Watson-Crick base
            pairing or does not disrupt normal Watson-Crick
30   base pairing, said protected thiol group being
            stable under conditions of chemical nucleic acid
            synthesis and/or conditions of enzymatic nucleic
            synthesis and being convertible to a reactive

thiol (SH) after said synthesis.

Preferred nucleotide analogs of the present invention are the protected phosphoramidites or 5' mono, di and triphosphates of modified cytosine or uridine bases for use in the chemical

5   synthesis of DNA and RNA by the phosphoramidite method or enzymatic synthesis with polymerases, ligases, or other nucleotide polymerizing enzymes.

In addition, the nucleotide analogs of the present invention include other nucleoside phosphates, containing 3',

10 5', or 3',5' monophosphates, diphosphates, or triphosphates and further optionally comprising 2' mono, di, or triphosphates when the nucleoside is a ribonucleoside.

Another aspect of this invention provides nucleic acids and oligonucleotides containing the subject nucleotide analogs

15 having a protected thiol group on a base moiety of that nucleic acid or oligonucleotide.   A method is also provided to synthesize these nucleic acids or oligonucleotides, deprotect the thiol group and attach a functional group to the reactive thiol moiety.

20   Yet another aspect of this invention is directed to a method of synthesizing the subject base-protected nucleotide analogs.

DESCRIPTION OF THE DRAWINGS

Figure 1: Synthetic pathway for the DNP-labeled phosphora-

25 midite (compound 7).   Reaction conditions are as follows a) SCNCl, CH_3CO_2H;   b) DTT, MeOH, EDTA;   c) 2,4-dinitrofluoro-benzene, Et_3N, CH_3CN;   d) NaOMe, MeOH;   e) $DMT^+BF_4^-$, DBMP, CH_3CN;   f) 2 - c y a n o e t h y l - N , N , N' , N' - t e t r a i s o p r o p y l - phosphorodiamidite, tetrazole, CH_3CN.

30   **Figure 2:** A. RP HPLC chromatograms of Oligonucleotides A and B. The purified 7-mers were analyzed by analytic HPLC on an Ultrasphere ODS column, eluted with a 15 minute linear gradient from 5% to 20% (v/v) acetonitrile in 20 mM TEAA (pH 7.1). (a) Oligonucleotide A, GTA TGT A, eluted at 9.8 min.   (b) DNP-

35 labeled oligonucleotide B, GTA T*GT A, eluted at 11.6 min, where T represents incorporation of the DNP-analog described in Figure 1.   B.RP   HPLC   profile   of   enzymatic   degradation   of

7

Oligonucleotides A and B.   The 7-mers were subjected to enzymatic treatment with Nuclease Pl and CIP.   The reaction mixture was analyzed by analytical HPLC on an Ultrasphere ODS column, eluted with a linear gradient of acetonitrile in 100 mM TEAA (pH 7.1) from 0% to 5% (v/v) in 5 min, 5% to 6% (v/v) in 15 min., 6% to 50% (v/v) in 5 min and finally 50% to 80% (v/v) in 15 min.   (a) oligonucleotide A reaction.   (b) DNP-labeled Oligonucleotide B.

Figure 3: A.: Gel electrophoretic analysis of DNP-labeled oligonucleotides. The[^{32}P] labeled oligonucleotides C-F (Table 1) were analyzed by electrophoresis on a 25% polyacrylamide-7M urea gel (acrylamide/methylene bisacrylamide=19:1).   Lane 1: Commercial 24-mer marker, CGC GGA ATT CTC ATG CAT TGC CCA (SEQ. ID. NO:5);   Lane 2; 24-mer, Oligonucleotide E;   Lane 3: DNP-labeled 24-mer, Oligonucleotide F;   Lane 4:   22-mer, Oligonucleotide C;   Lane 5: DNP-labeled 22-mer, Oligonucleotide D.   B.: Gel electrophoretic analysis of the deprotection and alkylation reactions.   Oligonucleotides E and F were reacted with 1.4 M BME. Half of these reactions were alkylated with 5-IAF.   The treated oligonucleotides were analyzed by electrophoresis on a 25% polyacrylamide-7M urea gel (acrylamide/ methylene bisacrylamide=19.1). Lane 1: Oligonucleotide E + 1.4 M BME at room temperature;   Lane 2: Lane 1 reaction + 5-IAF; Lane 3: Oligonucleotide E + 1.4 M BME at 45°C;   Lane 4: Lane 3 reaction + 5-IAF;   Lane 5: Oligonucleotide F + 1.4 M BME at room temperature;   Lane 6: Lane 5 reaction + 5-IAF;   Lane 7: Oligonucleotide F + 1.4 M BME at 45°C;   Lane 8: Lane 7 reaction + 5-IAF.

Figure 4: Analysis of PCR products.

A. PCR products were analyzed by electrophoresis on a 1.5% agarose gel. DNA was detected with ethidium bromide. Lane 1:λ Hind III markers;   Lane 2: PCR product using Oligonucleotide E as a primer;   Lane 3: PCR product using DNP-labeled Oligonucleotide F as a primer. XC represents xylene cyanol and BPB represents bromophenol blue.

B. Dot Blot immunodetection of DNP-labeled DNA.   PCR products were attached to a nylon membrane, incubated with a primary antibody to DNP and a 2° antibody linked to horseradish

peroxidase. Chemiluminescence was used for detection. Upper panel: PCR product using DNP-labeled Oligonucleotide F as a primer; Lower panel: PCR product using Oligonucleotide E as a primer.

5   **Figure 5:**   Protein Crosslinking to Site-Specifically Modified DNA

To analyze the DNA-protein complexes in the transcription complex a deoxynucleoside phosphoramidite was incorporated site-specifically into a single-stranded DNA oligonucleotide via
10 automated synthesis. In these studies, the nucleotide analog was modified with a masked reactive thiol group on the non-basepairing 5 position of deoxyuridine (5-S-DNP-dU, Figure 1). The oligonucleotide was then radioactively labeled uniquely on the 5' phosphate of the nucleotide analog with polynucleotide
15 kinase and [γ-^{32}P] ATP (Step 1). The radioactively labeled, base-modified oligonucleotide was then annealed to a single-stranded DNA template (Step 2) and incorporated site-specifically into double stranded DNA by primer extension (Step 3). The protecting group was then removed from the double-
20 stranded DNA by treatment with β-mercaptoethanol (Step 4), and after removal of the BME, a functional group was added to the unmasked thiol group. For crosslinking, the DNA was reacted with azidophenacyl bromide to attach a photoreactive azide group (Step 4). RNA polymerase and transcription factors are then
25 added to assemble a complex on the modified DNA, and the complex is irradiated with ultraviolet light to covalently join the DNA to proteins which are bound adjacent to the crosslinker (Step 5). Lastly, the DNA-protein complex is treated with a nuclease which leaves 5' NMPs to digest the DNA to a small fragment (Step
30 6). This will leave a single radioactive label on the 5' phosphate of the nucleotide involved in the DNA-protein crosslink.
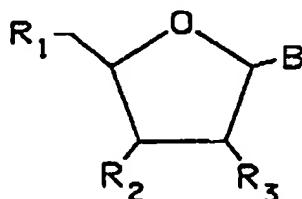
DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a series of novel
35 nucleotide analogs which are masked synthons for use as intermediates in chemical or enzymatic synthesis of nucleic acids, including synthesis of both oligonucleotides and

9

polynucleotides.   The nucleotide analogs of this invention, which contain a protected thiol group, can thus be incorporated into DNA or RNA under standard synthetic conditions without loss of the thiol protecting group.   This stability of the thiol

5   protecting group permits site-selective introduction into a nucleic acid of the nucleotide analog in a manner which facilitates later addition of a functional group at that site. Thus, the subject oligonucleotides (or polynucleo-tides) can contain one (or more) of the subject nucleotide analogs.

10   Thus, functional groups which can not withstand the conditions for chemical nucleic acid synthesis, especially during automated synthesis, or which may be too bulky or sterically hindered for enzymatic nucleic acid synthesis can be readily incorporated into the final oligonucleotide (or

15   polynucleotide) product.   Such derivatized nucleic acids have great utility in studying protein-nucleic acid or nucleic acid-nucleic acid interactions, as well as the potential for diagnostics and delivery of therapeutics.   Hence, the present invention permits the skilled artisan to place chemical tags

20   such   as   crosslinking   groups,   fluorescent   molecules, radioisotopes or other reporter molecules at specific positions on nucleic acid molecules for analysis of molecular mechanisms, for creation of diagnostic probes, for therapeutics, for antisense therapeutics and for many other purposes.

25   In particular, the nucleotide analogs of the present invention are intermediates for the chemical synthesis of DNA and RNA by manual or automated techniques and are represented by the formula:

wherein

30   R_1 is -H, -OH, a mono, di, or triphosphate group, or
        $-OR_4$;

     R_2 is -H, -OH, a mono, di, or triphosphate group, a
        phosphoramidite   group,   a   phosphorothioamidite

SUBSTITUTE SHEET (RULE 26)

group, a phosphonate group, an O-substituted monophosphate group, $-OR_4$, or a solid support bonded via an O at the 3' position;

R_3 is -H, -OH, a mono, di, or triphosphate group, or $-OR_4$;

R_4 is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at a position on said base, preferably the 5 position, that is not involved in Watson-Crick base pairing or does not disrupt normal Watson-Crick base pairing, said protected thiol group being stable under conditions of chemical nucleic acid synthesis and/or conditions of enzymatic nucleic synthesis and being convertible to a reactive thiol after said synthesis.

As used herein, protecting groups defined by R_4 are known to those of ordinary skill in the art and include any known protecting group suitable for protection of the 2', 3' or 5' hydroxyls of ribose and the 3' or 5' hydroxyls of deoxyribose sugars. In this regard, Greene et al. (1990) Protecting Groups in Organic Synthesis, 2nd ed., John Wiley & Sons, Inc., New York, NY, provides a comprehensive review of protecting groups, and methods of preparing the corresponding protected compounds, which can be used for different reactive groups, including reactive hydroxyl groups.
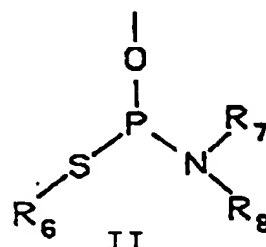
Accordingly, examples of protecting groups defined by R_1, R_2 and R_4 are lower alkyl, lower cyanoalkyl, lower alkanoyl, aroyl, aryloxy, aryloxy-lower alkanoyl, haloaryl, fluorenylmethyloxy-carbonyl (FMOC), levuloyl, 9-phenylxanthene-9-yl, trityl, p-monomethoxytrityl (MMTr), p-dimethoxytrityl (DMTr), isopropyl, isobutyl, 2-cyanoethyl, acetyl, benzoyl, phenoxyacotyl, halophenyl, 1-(2-chloro-4-methylphenyl)-4-methoxy-4-piperidinyl, 2'-acetal, O-nitrobenzyl, tert-butyldimethylsilyl (TBDMS), tetrahydrofuranyl, 4-methoxytetrahydropyranyl and related groups.

Preferred R_1 protecting groups include 5'OH protecting groups especially DMTr, MMTr, FMOC, levuloyl and 9-

11

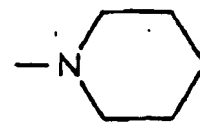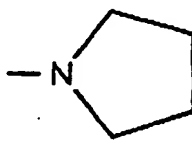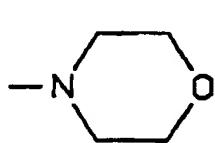phenylxanthene-9-yl groups, and most preferably DMTr. Preferred R_2 protecting groups include 3'OH protecting groups, especially the acetyl group. Preferred R_4 protecting groups include 2'OH protecting groups, especially 2'-acetal, tetrahydrofuranyl, 4-methoxytetrahydropyranyl and 1-(2-chloro-4-methylphenyl)-4-methoxy-4-piperdinyl groups, and most preferably TBDMS.

As defined herein, solid supports include controlled pore glass (CPG), polystyrene silica, cellulose, nylon, and the like. Preferred solid supports are CPG and polystyrene. An especially preferred solid support is CPG.

R_2 can be taken with the oxygen atom to which it is attached, to form a phosphoramidite, phosphorothioamidite, a phosphonate, an O-substituted monophosphate, or any other group compatible with chemical nucleic acid synthesis, especially automated DNA or RNA synthesis. As used herein, the phosphoramidite and phosphorothioamidite groups have the general formulas I and II, respectively:

wherein R_6 is lower alkyl, n-cyano alkyl, substituted lower alkyl, aryl, aralkyl, substituted aralkyl and the like, and R_7 and R_8 are independently lower alkyl or when taken together with the nitrogen to which they are attached comprise cyclic groups including:

preferably the R_6 group of the phosphoramidite group (Formula I) is 2-cyanoethyl (CED) or methyl. Accordingly, the preferred phosphoramidites of this invention are also referred to as CED phosphoramidites or methyl phosphoramidites.
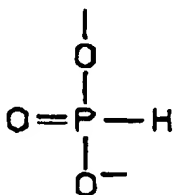
12

As used herein, the term lower alkyl, when used singly or in combination, refers to alkyl groups containing one to six carbon atoms. Lower alkyl groups can be straight chain or branched and include such groups as methyl, ethyl, propyl, isopropyl, n-butyl, sec-butyl, isobutyl, t-butyl, pentyl, isopentyl, neopentyl, hexyl and the like. The preferred alkyl groups contain one to four carbon atoms.

The term aryl, when used alone or in combination, refers to an aromatic ring containing six to ten ring carbon atoms. The aryl group includes phenyl, and 1- or 2-naphthyl. The preferred aryl group is phenyl.

The term aralkyl refers to aryl groups as described above to which substituents are attached to the aryl by an alkylene bridge. The most preferred aralkyl group is benzyl.

When R_7 and R_8 are lower alkyl, the preferred groups are each isopropyl groups.

As used herein a phosphonate group (or H-phosphonate) is represented by the formula:

$$O=\overset{\displaystyle O}{\underset{\displaystyle O^-}{\overset{\displaystyle |}{\underset{\displaystyle |}{P}}}}-H$$

and may also be provided as salts, and preferably as triethylammonium salts.

As used herein, O-substituted monophosphates have the formula:

$$R_9-O-\overset{\displaystyle O}{\underset{\displaystyle O^-}{\overset{\displaystyle |}{\underset{\displaystyle |}{P}}}}=O$$

wherein R_9 is lower alkyl, haloalkyl, aryl, haloaryl, or heteroaromatic. By haloalkyl or haloaryl is meant alkyl or aryl groups, respectively, which have been substituted with one or more halogen atoms including F, Cl, Br, or I. Preferred halo

substituents are Cl and Br.   Preferred R_4 groups include 2-chlorophenyl, 2,5-dichlorophenyl, 2,2,2-trichloroethyl and the N-oxide of 4-methoxypyridine-2-methylene groups.

Modified pyrimidine bases have a protected thiol group attached at a position on the base which is not involved with Watson-Crick base pairing or which does not disrupt normal Watson-Crick base pairing. The protected thiol groups of this invention are stable, i.e. not removable, under the conditions used for chemical synthesis of DNA or RNA, and particularly, under the conditions employed in automated DNA or RNA synthesis. Furthermore, the protecting group of the thiol is removable under conditions which do not disrupt the integrity of the oligonucleo-tide or polynucleotide.   In other words, after a nucleotide analog of the present invention has been incorporated into an oligonucleotide, for example, the protected thiol group can be converted to a reactive thiol (SH) to which functional groups can subsequently be added using thiol-modifying reagents.

In accordance with this invention the protected thiol groups include, but are not limited to, the dinitrophenyl group. As used herein, Watson-Crick base pairing refers to the hydrogen bonding pattern of adenine-thymine (AT) base pairs, adenine-uracil (AU) base pairs, or of guanine-cytosine (GC) base pairs. Accordingly, the preferred protected thiol position on pyrimidines is the 5 position on the pyrimidine (C,U) ring.

The preferred bases of the present invention are thus pyrimidines such as cytosine (C), and uracil (U). Cytosine bases contain an additional reactive group, specifically an exocyclic amine, which must also be protected during assembly of the nucleotide chain via chemical synthesis. Accordingly, bases of this invention can have additional protecting groups attached as needed to any of the ring positions. Such protected bases are well known in the art and include but are not limited to, N^2, N^4-anisoyl cytosine, N^4-benzoyl cytosine, and N^4-isobutyryl cytosine. The N^4-isobutyryl cytosine is known by the trade name FOD[TM] base protection (Applied Biosystems, Inc.). Use of these further protected bases is compatible with incorporation of the protected thiol groups as well as subsequent deprotection reactions.

The bases of this invention also include any related base analog that is capable of base pairing with a guanine or adenine, the corresponding protected analogs as set forth above for use in chemical synthetic methods to produce DNA and RNA. For example, such base analogs include, but are not limited to, pseudocytosine, isopseudocytosine, 4-acetylcytosine, 2'-O-methylcytosine, dihydro-uracil, 2'-O-methyluracil, 2'-O-methyl-pseudouracil, 1-methyl-pseudouracil, 3-methylcytosine. Bases attached to a ribose or deoxyribose sugar in an α-anomeric configuration can also be present.

In a preferred embodiment the bases of the invention are C, U, and N^4 protected C.

The compounds of the present invention also include nucleoside phosphates of the formula:

wherein:

R$_1$ is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R$_2$ is -H, -OH, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, -OR$_4$, or a solid support bonded via an O at the 3' position;

R$_3$ is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R$_4$ is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at the 5 position on said base.

In the case of the modified bases the protected thiol group is also stable, i.e. not removable, under the conditions used for enzymatic synthesis of DNA or RNA. Additionally, the pyrimidine bases generally do not require protection of the

additional reactive groups such as the exocyclic amines for enzymatic incorporation into DNA or RNA. Accordingly, the preferred bases for this class of compounds includes C and U.

5   In accordance with this invention the phosphate groups of R_1, R_2 and R_3 embody all the phosphorylated forms at the C_3' and C_5' positions of the sugar moiety (and C_2' position when the sugar is ribose) and include monophosphates, diphosphates, triphosphates and tetraphosphates. Preferably the phosphate group is a triphosphate for R_1.

10   The preferred compounds of this invention include the phosphoramidites and 5'triphosphates of 5-SDNP-dU, 5-SDNP-U, 5-DNP-dC, and 5-SDNP-C. In the case of the phosphoramidites, the base moieties can contain additional protecting groups on the exocyclic amines. The preferred compounds also include the
15   5'triphosphates of 5-S-S-Et-dU, 5-S-S-Et-U, 5-S-S-Et-dC and 5-S-S-Et-C, 5-S-CH_2-$NHCOCH_2Ph$-dU, 5-S-CH_2-$NHCOCH_2Ph$-U, 5-S-CH_2-$NHCOCH_2Ph$-C, 5-S-CH_2-$NHCOCH_2Ph$-dC, 5-SCH_2CH_2DNP-dU, 5-SCH_2CH_2DNP-U, 5-SCH_2CH_2DNP-dC, and 5-SCH_2CH_2DNP-C.

The nucleotide analogs of the present invention can be
20   prepared by adding a protected thiol group to the base moiety of the desired nucleoside. The so-modified nucleoside can then be phosphorylated to produce a nucleotide analog phosphate compound of this invention using conventional phosphorylation techniques. To produce a phosphoramidite, phosphonate, phosphorothioamidite
25   or O-substituted monophosphate of this invention, the 5'-OH or 2'-OH groups of above-modified nucleoside are protected as necessary by addition of the desired protecting group by standard methodology. This protection step(s) is (are) followed by conversion to the nucleotide-analog phosphoramidite,
30   phosphorothioamidite, phosphonate or O-substituted monophosphate by reaction of the 3'OH of the nucleoside with the appropriate modifying group. Similarly, the 3'OH can be attached to a solid support, such as CPG, or another protecting group using conventional methodology available to the ordinarily skilled
35   artisan in this field.

For example, to prepare protected thiol-groups at the 5 position of uridine or 2'-deoxyuridine nucleosides, the sulfur atom must be incorporated on the base, for which the method of

SUBSTITUTE SHEET (RULE 26)

16

Nagamachi et al. (1972, J.C.S. Chem. Commun. 18:1025-6) has been modified. An equivalent amount of potassium thiocyanate is dissolved into a solution of chlorine gas in glacial acetic acid to produce the thiocyanogen chloride in solution. An iodometric
5   titration is performed with Cl_2/HOAc to determine the amount of Cl_2 present. The solution can be used directly or filtered prior to the next step. In either case, the acetyl-protected U or dU is added all at once with stirring. This reaction is maintained with stirring at room temperature for about 1.5 hours until the
10 reaction is completed. An excess of cyclohexene or other quenching agent is then added to quench any remaining thiocyanogen chloride. Quenching is complete in 15 to 60 min, and usually in about 30 min. After removing solvents and organic residue in vacuo, the remaining residue can be purified
15 by silica gel chromatography with a gradient of petroleum ether and ethyl acetate or with other chromatographic methods. These acetyl-protected 5-thiocyanato dU and U nucleosides can then be reduced to 5-mercapto nucleosides for further derivatization as described herein to produce the nucleotide analogs of the
20 present invention.

The mono-, di-, tri or tetra phosphate compounds can be prepared by a similar reaction by beginning with the appropriate starting 5-thioprotected nucleoside. If necessary, the 2'-OH of the ribose sugar can be protected prior to the reaction using
25 any of the known 2'-OH protecting groups by conventional techniques.

The described synthesis and utilization of 5-thiocyanato-deoxyuridine phosphoramidite for placement of the sulfur directly on the base has been previously described (Bradley &
30 Hanna, 1992). However, the synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-S-(2,4-dinitrophenyl) mercapto-2'-deoxyuridine-3'-O-(2-cyanoethyl-N,N'-diisopropyl) phosphoramidite is more reproducible and the product is more stable than 5-thiocyanatodeoxyuridine phosphoramidite and
35 represents a more desirable analog. A variety of functional groups can be attached to the reactive thiol. Once this modified nucleotide is incorporated into an oligonucleotide or nucleic acid, the protecting group can be removed to unmask the

SUBSTITUTE SHEET (RULE 26)

reactive thiol, making it available for further derivitization
with thiol-specific agents. This method has the advantage that
an oligonucleotide tagged with this analog can be synthesized
and stored for long periods without removal of the protecting
5   group. When needed, an aliquot can then be deprotected and
modified. In addition, in some cases the protecting groups are
themselves antigenic (e.g., ⁻DNP) and may be retained in the
nucleic acid product for detection utilizing antibodies.

One method used to produce one of the nucleotide analog
10   phosphoramidites of this invention and described in detail in
Example 1 below is shown in Fig. 1. Briefly, as depicted, 5-
SDNP-dU or 5-SDNP-U are prepared by reduction of the SCN,
protection of the thiol as the DNP ether, and de-esterification
of the sugar ring. The protected-thiol nucleoside thus formed
15   is then reacted with the DMT⁺BF₄⁻, or another 5'OH protecting
group, in an anhydrous organic solvent in the presence of an
organic base for 2-24 hours until the reaction is complete. The
resulting product can be isolated by liquid chromatographic
methods and then converted to a CED phosphoramidite by reaction
20   with 2-cyanoethyl N, N, N', N'-tetraisopropyl phosphordiamidite
and tetrazole under anhydrous conditions and in an inert
atmosphere. This reaction is preferably stirred for 1 hour,
although this time can be varied, the solvents removed _in vacuo_
and the residue purified by chromatographic methods. Exposure
25   to atmosphere is acceptable but should be minimized. This
phosphoramidite can be stored dry under positive pressure of
inert atmosphere (argon or nitrogen) in a tightly sealed
container at low temperature.

Other phosphoramidites of this invention can be prepared by
30   reaction with the appropriate chlorophosphoroamidite. For
example, the O-methylphosphoramidites can be prepared by
reacting the 5-SDNP-5'-DMTr-dU with N,N-
diisopropylmethylphosphonamidic chloride by conventional
techniques. Similarly, the phosphorothioamidite, phosphonates,
35   O-substituted monophosphate can be prepared using conventional
techniques with commercially available reagents. As mentioned
above to avoid unwanted side reactions, protection of the 5'OH
and 2'OH groups on the sugar moiety and the exocyclic amine

groups on the base moiety may be necessary before the final reaction step which produces the nucleotide analogs of this invention.

In general the chemical synthetic routes to nucleotide phosphoramidites, phosphorothioamidites, phosphonates and O-substituted monophosphates, as well as nucleotide phosphates, are well known. In addition, nucleoside phosphates can be enzymatically synthesized. Chemical synthetic techniques for these compounds, as well as the common synthetic routes to prepare RNA and DNA, have been described in many sources. Particularly, useful references include Gait (1984) Oligonucleotide Synthesis: A Practical Approach, IRL Press, Oxford; Blackburn et al. (1990) Nucleic Acids in Chemistry and Biology, IRL Press, Oxford, especially Chap. 3; Chaps 13-16 of Methods in Enzymology, Vol. 154 (Wu et al., eds.) Academic Press, San Diego, CA, 1987; Chaps. 13-14 in Methods in Molecular Biology, Vol. 4 (Walker, ed.), Humana Press, Clifton, NJ, 1988; and Uhlmann et al. (1990) Chem. Rev. 90:544-584.

In addition to providing methods for chemical synthesis of DNA and RNA, some of these references (particularly Gait and Uhlmann et al.) describe the reactions and methodology for adding protecting groups to 5'OH, 3'OH and 2'OH groups, and exocyclic amine groups (N^6 of adenine, N^4 of cytosine, N^2 of guanine). These references also provide information and protocols to attach nucleotides to solid supports which protocols are useful for attaching the base-protected nucleotide analogs of this invention.

The synthesis of 5-SDNP-dC and 5-SDNP-C can be accomplished by reaction of 5-Br-dC or 5-Br-C with sodium hydrogen sulfide, as described by Solan, followed by conversion into the DNP thiol ether. In these cases, it may be necessary to first protect the 4 amino group of cytosine (e.g., as the iso-butoxy amide) and complete the remaining protection steps (5'-O-DMTr and 2'-O-TBDMS, if needed) before preparation of a CED phosphoramidite or conversion to the triphosphate (Sung, 1982, J. Org. Chem. 47:3623). Alternatively, 5-S-dU or 5-S-U can be converted to the corresponding compounds. Several methods are available for this conversion of U to C, including Sung, Xu, and MacMillan.

19

The conversion of 5-SDNP-dU and 5-SDNP-U to 5'phosphates, phosphoramidites, phosphorothioamidites, phosphonates, and O-substituted monophosphates has been described herein above. All of these reaction schemes can be used to produce the
5   corresponding thiol protected cytosine analogs. If necessary, various protecting groups for the 5'OH, 3'OH or 2'OH groups as well as the exocyclic amines can be added in accordance with the methodology described herein.

Another aspect of this invention relates to the
10 oligonucleotides or polynucleotides containing the nucleotide analogs of this invention and a method of preparing such nucleic acids using the subject nucleotide analogs. Oligonucleotides and polynucleotides of this invention are made by standard methods of chemical (automated or manual) synthesis or enzymatic
15 synthesis of DNA and RNA. Such methods are well known in the art. In chemical synthesis, the nucleotide analog of this invention is substituted for a particular nucleotide at the desired point in the synthesis.

After incorporation of the nucleotide analog and complete
20 synthesis of the oligonucleotide or polynucleotide, the thiol can be deprotected and reacted with any number of thiol-modifying reagents to attach a functional group at that point on the oligonucleotide or polynucleotide. Deprotection of S-DNP or S-S-Et can be accomplished by treatment with β-mercaptoethanol
25 or by other means of reducing sulfides. In a preferred method for the DNP analog, deprotection is accomplished by treating the oligonucleotide with 1.4M BME at 45°C for 4 hours or at room temperature overnight.

Any variety of functional groups can then be added to the
30 reactive thiols group generated by the deprotection step. Such functional groups include cross-linking groups, photoactive cross-linking groups (e.g. arylazides), and reporter molecules such as radioisotopes, biotin, enzymes and fluorescent markers. The methods for adding functional groups are well know in the
35 art.

Yet another aspect of this invention provides a method of preparing the nucleotide analogs of the invention. In particular, this method involves preparing a thiol protected

SUBSTITUTE SHEET (RULE 26)

20

nucleoside or nucleotide base wherein said thiol is attached to a position on said base that is not involved in Watson-Crick base pairing or does not disrupt normal Watson-Crick base pairing; reacting said nucleoside or nucleotide base under
5   conditions to effect conversion of said base to a phosphoramidite, phosphorothioamidite, phosphonate, O-substituted monophosphate or phosphate nucleotide analog and under conditions which do not destroy the protected thiol; and recovering said analog. In accordance with this invention, this
10   method is accomplished as described above for synthesis of the subject nucleotide analogs. Recovery of the analogs can be accomplished by HPLC, FPLC or other chromatographic separation techniques.

       A further aspect of this invention provides a method of
15   synthesizing a nucleic acid with a functional group by incorporating a thiol protected nucleotide analog in accordance with this invention into a nucleic acid by a chemical or enzymatic method for nucleic acid synthesis; recovering the nucleic acid containing the analog; deprotecting the analog of
20   that nucleic acid to produce a nucleic acid containing a reactive thiol group; reacting the reactive thiol group with a thiol-modifying reagent to thereby attach a functional group and produce the nucleic acid with the functional group; and recovering the nucleic acid with the functional group. In
25   accordance with this invention, this method is accomplished as described hereinabove for synthesis of oligonucleotides and polynucleotides (See for example Gait or Sambrook et al). As used herein, nucleic acids include oligonucleotides and polynucleotides. Preferably, the oligo-nucleotides range in
30   size from about 5 to about 100 nucleotides. Polynucleotides range in size from 100 nucleotides to 10 kb or more. Recovery of the analogs can be accomplished by HPLC, FPLC, other chromatographic techniques, extraction, phase separation or precipitation.

35   While the invention will now be described in connection with certain preferred embodiments in the following examples so that aspects thereof may be more fully understood and appreciated, it is not intended to limit the invention to these

particular embodiments.   On the contrary, it is intended to cover all alternatives, modifications and equivalents as may be included within the scope of the invention as defined by the appended claims.   Thus the following examples, which include
5   preferred embodiments, will serve to illustrate the practice of this invention, it being understood that the particulars shown are by way of example and for purposes of illustrative discussion of preferred embodiments of the present invention only and are presented in the cause of providing what is
10   believed. to be the most useful and readily understood description of formulation procedures as well as of the principles and conceptual aspects of the invention.

                              EXAMPLES

Procedures
15   **Materials.** Unless otherwise stated, starting materials for the chemical synthesis of the phosphoramidite were obtained from Aldrich, Sigma, Fluka or Fisher Scientific and were used without further purification.   Anhydrous HOAc was prepared by fractional freezing (Nagamachi et al., 1974).   Anhydrous CH_2Cl_2 was
20   distilled from phosphorus pentoxide.   Anhydrous CH_3CN was purchased from Cruachem.   Dowex 50W x4 was purchased from BioRad.   Analytical thin layer chromatography was performed on Whatman silica gel 60 plates with fluorescent indicator. Column chromatography was performed with silica gel 60A from American
25   Scientific Products. Reagents for the oligonucleotide syntheses were obtained from Cruachem.   Standard exocyclic amine protecting groups were used: Benzoyl for dA and dC and isobutyryl for dG.   Solvents utilized for the reverse phase chromatographic purification of the oligonucleotides were of
30   HPLC grade.   The fluorescent thiol modifying agent, 5-IAF was purchased from Molecular Probes (Eugene, Oregon).
         Buffers.   The buffers used were as follows: buffer A, 100 mM TEAA (pH 7.1); buffer B, 20 mM TEAA (pH 7.1); buffer C, 50 mM Tris-HCl (pH 4.9), 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$; buffer D, 50 mM
35   Tris-HCl (pH 9.0), 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 0.5 mM spermidine; buffer E, 7 M urea, 0.05% (w/v) xylene cyanol; buffer F, 5% (v/v) glycerol, 0.04% (w/v) xylene cyanol, 0.04% (w/v)

22

bromophenol blue.

Enzymes and Antibodies. Nuclease P1 was obtained from Pharmacia Biotech. CIP was supplied by Promega. T4 Polynucleotide Kinase was purchased from New England Biolabs. AmpliTaq Polymerase was obtained from Perkin Elmer. Rabbit polyclonal anti-DNP antibody was purchased from Sigma. Anti-rabbit Ig, horseradish peroxidase-linked whole antibody was obtained from Amersham.

Analytical Methods. Melting points were determined on a Mel-Temp melting point apparatus and are uncorrected. ^{1}H NMR and ^{31}P NMR spectra were recorded on a Varian XL-300 or 500 spectrometer, respectively. Chemical shifts are reported in parts per million (δ) relative to internal tetramethylsilane or to external 85% phosphoric acid. UV spectra were measured on either a LKB Biochem Ultraspec II or a Beckman DU 7500 spectrophotometer. IR spectra were recorded with a Perkin Elmer 710B Infrared spectrometer. FAB MS were recorded on a VG analytical ZAB-E spectrometer. Unless otherwise stated, all chemical reactions were performed under a nitrogen atmosphere.

HPLC analyses were performed with a Beckman LC system which was equipped with a 126 solvent delivery module and a 168 diode array detector and was controlled with Beckman System Gold Software. A reverse phase C18 column (Beckman Ultrasphere ODS, 250 x 4.6 mm, i.d.) and guard column (Upchurch, 1 cm x 4.3 mm, i.d.) were utilized with both analytical and semi-preparative separations. Sample loops of 1 mL and 50 μL were used for semi-preparative and analytical analyses, respectively. All HPLC analyses were performed using a gradient solvent system composed of triethylammonium buffers and acetonitrile at a flow rate of 1 mL/min. Unless otherwise stated, UV absorption was monitored at a wavelength of 260 nm.

Abbreviations: DNP,2,4-dinitrophenyl; 5-IAF, 5-iodoacetamido-fluorescein; CIP, Calf intestinal alkaline phosphatase; TEAA, triethylammonium acetate; TEAB, triethylammonium bicarbonate; DMT, dimethoxytrityl; RP HPLC, reverse phase high pressure liquid chromatography; PCR, polymerase chain reaction; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; DMTCl, dimethoxytrityl

23

chloride; DMT$^+$BF$_4$, dimethoxytrityl tetrafluoroborate; BME, β-mercaptoethanol; HR FAB MS, high resolution fast atom bombardment mass spectroscopy.

Example 1.   Preparation of 5-SDNP-dU-phosphoramidite

5   Shown in Figure 1 and described in detail below is the pathway used for the synthesis of the DNP-labeled phosphoramidite, Compound 7.

Synthesis of 3',5'-0-Diacetyl-5-thiocyanato-2'-deoxyuridine (Compound 2).

10   Anhydrous HOAc (150 mL) was added to a three-neck round bottom flask equipped with a vacuum adapter and teflon needle. Dry Cl$_2$ was slowly bubbled through the teflon needle into the reaction vessel for a brief period. A 5 mL aliquot of the Cl$_2$/HOAc solution was removed and subjected to iodometric titration to

15   determine the amount of Cl$_2$ added (2.9 g, 41 mmol) (Flaschka et al., 1969). KSCN, previously dried overnight at 100°C, (4.4 g, 45 mmol) was added to the reaction vessel. The reaction mixture was allowed to stir at room temperature for 30 min. prior to the addition of 3',5'-0-diacetyl-2'-deoxyuridine, 3',5'-O-Diacetyl-

20   2'-deoxyuridine was prepared by reaction of deoxyuridine with acetic anhydride in pyridine for an 82% yield: mp=106-108°C(Lin & Gao, 1983) (1.1 g, 3.5 mmol). After the mixture was stirred for an additional 1.5 hours, cyclohexene (10 mL) was added. The mixture was then stirred for 15 min. and filtered; the filtrate

25   was concentrated under vacuum. Residual HOAc was removed azeo-tropically with toluene and the residue was triturated with petroleum ether. The petroleum ether was decanted and the residue subjected to column chromatographic purification [petroleum ether: EtOAc(2:1 - 1:3)] which yielded a semi-solid.

30   Recrystallization from EtOH:EtOAc provided Compound 2 as a white solid (1.05 g; 82%):mp 154-156°C; IR(KBr) 2140cm^{-1} (SCN stretch); ^{1}H NMR (CDCl$_3$) δ 9.12(bs, 1H, N$_3$H), 8.18 (s, 1H, H$_6$), 6.14 (dd, J=7.8, 5.7 Hz, 1H, H$_{1'}$), 5.13 (m, 1H, H$_{3'}$), 4.24 (m, 1H, H$_{5'}$), 4.19-4.15 (m, 2H, H$_{4'}$ and H$_{5'}$), 2.32 (ddd, J= 14.4, 5.7, 2.1 Hz, 1H,

35   H$_{2'}$), 2.13-2.07 (m, 7H, 2 CH$_3$ and H$_{2'}$); UV (MeOH) 270 nm.

SUBSTITUTE SHEET (RULE 26)

Synthesis   of   3',5'-O-Diacetyl-5-mercapto-2'-deoxyuridine, (Compound 3).

This reaction should be performed in a fume hood.   EDTA (25 mL; 0.1 M, pH 7.7) and DTT (293 mg, 1.90 mmol) were added to a solution of thiocyanate 2 (207 mg, 0.561 mmol) in MeOH (30 mL). The reaction mixture was allowed to stir at room temperature for 35 min. and was then filtered.   The filtrate was reduced to half by rotary evaporation, made acidic with 10% (v/v) H_2SO_4 and extracted with CH_2Cl_2.   The CH_2Cl_2 extract was concentrated under reduced pressure and the residue was triturated with H_2O to provide Compound 3 as a white solid (148 mg; 77%).   TLC analysis showed a yellow spot upon development with 2,2'-dithiobis(5-nitropyridine) (Grassetti & Murray, 1969): mp=159-163°C; IR(KBr) 2530 cm^{-1} (SH stretch); ^1H NMR (CDCl$_3$) δ 8.70 (bs, 1H, N$_3$H), 7.67 (s, 1H, H$_6$), 6.30 (dd, J=8.2, 5.8 Hz, 1H, H$_{1'}$), 5.22 (m, 1H, H$_{3'}$), 4.43-4.25 (m, 3H, H$_{4'}$ and H$_{5'}$), 3.60 (s, 1H, SH), 2.50 (ddd, J=14.4, 5.8, 2.2 Hz, 1H, H$_{2'}$), 2.21-2.10 (m, 7H, 2 CH$_3$ and H$_{2'}$); UV (0.1 M EDTA w/DTT, pH 7.76) 260, 334 nm.   (The broad melting point range may indicate the presence of 5,5"-dithiobis(3',5'-diacetyldeoxyuridine).   On the average, 7% disulfide analog was present in the sample as determined by UV absorbance (λmax 335 nm) in the presence and absence of DTT (Bardos & Kalman, 1966)).

Synthesis of 3',5'-O-Diacetyl-5-S-(2,4-dinitrophenyl)mercapto-2'-deoxyuridine, (Compound 4).

2,4-Dinitrofluorobenzene (57 mg, 0.31 mmol) and Et$_3$N (200 μL) were added to a solution of mercaptan 3 (86 mg, 0.25 mmol) in anhydrous CH$_3$CN (10 mL).   The reaction mixture was allowed to stir at room temperature for 1.5 hours before the solvent was removed by rotary evaporation.   The crude material was purified using column chromatography [Hexanes: EtOAc(4:1 to 1:1)].   The product was obtained as a yellow solid (113 mg, 89%): IR(KBr) 1530, 1340 cm^{-1} (NO$_2$ stretch); ^1H NMR (CDCl$_3$) δ 9.14 (d,J=2.4 Hz, 1H, H$_3$ of DNP), 8.40 (bs, 1H, N$_3$H), 8.31 (dd, J=9.1, 2.5 Hz, 1H, H$_{5'}$ of DNP), 8.23 (s, 1H, H$_6$), 7.33 (d,J=9 Hz, 1H, H$_6$ of DNP), 6.32 (dd, J=7.8, 5.7 Hz, 1H, H1'), 5.25, (m, 1H, H$_{3'}$), 4.44-4.25 (m, 3H, H$_{4'}$ and H$_{5'}$), 2.55 (m, 1H, H$_{2'}$), 2.27 (m, 1H, H$_{2'}$), 2.13 (s, 3H, CH$_3$), 2.10 (s, 3H, CH$_3$); UV (MeOH) 264, 323 nm.

25

Synthesis of 5-S-(2,4-Dinitrophenyl)mercapto-2'-deoxyuridine, (Compound 5).

NaOMe (25% solution in MeOH; 82 μL; 0.36 mmol) was added to a solution of Compound 4 (92 mg, 0.18 mmol) in anhydrous MeOH (6
5   mL) and the reaction mixture was maintained at room temperature. After 2 hours, the mixture was treated with Dowex 50W x4 and filtered. The filtrate was reduced by rotary evaporation. The product was obtained as a yellow solid following column chromatographic purification of the reaction mixture [EtOAc:
10   petroleum ether (1:4 to 1:0)] (62 mg; 81%): IR(KBr) 1530, 1340 cm^{-1} (NO$_2$ stretch); ^{1}H NMR (Cd$_3$OD) δ 9.08 (d, J = 2.4 Hz, 1H, H$_3$ of DNP), 8.85 (s, 1H, H$_6$), 8.39 (dd, J = 9, 2.4 Hz, 1H, H$_5$ of DNP), 7.64 (d, J = 9 Hz, 1H, H$_6$ of DNP), 6.33 (t, J = 6.3 Hz, 1H, H$_{1'}$), 4.45 (m, 1H, H$_{3'}$), 4.00 (m, 1H, H$_{4'}$), 3.84 (dd, J – 12, 3 Hz, 1H,
15   H$_{5'}$), 3.75 (dd, J = 12, 3 Hz, 1H, H$_{5'}$), 2.50-2.33 (m, 2H, H$_{2'}$); UV (0.1 M TEAA, pH 7.1 with 57% CH$_3$CN) 262, 334 nm.

Synthesis of 5'-O-Dimethoxytrityl-5-S-(2,4-dinitrophenyl)mercapto-2'-deoxyuridine (Compound 6).

Anhydrous CH$_3$CN (6 mL) was added to a mixture of compound 5 (42
20   mg, 0.10 mmol), DMT$^+$BF$_4^-$(57 mg, 0.14 mmol) and 2,6-di-tert-butyl-4-methylpyridine (DBMP) (35 mg, 0.18 mmol). The resulting red solution was heated to reflux. After 1.5 hours, a second portion of DMT$^+$BF$_4$– (26 mg, 0.064 mmol) and DBMP (13 mg, 0.067 mmol) was added and the reaction mixture was maintained at
25   reflux for an additional 8 hours. The CH$_3$CN was removed by rotary evaporation. CH$_2$Cl$_2$ and 5% aq. NaHCO$_3$ were added to the residue. The mixture was stirred briefly and the resulting two layers separated. The NaHCO$_3$ layer was re-extracted twice with additional CH$_2$Cl$_2$. The organic layers were combined, dried
30   (Na$_2$SO$_4$) and reduced. The product was purified using column chromatography [CH$_2$Cl$_2$:MeOH:pyridine (99.5:0:0.5 to 98.5:2:0.5)] (27 mg: 37%); ^{1}H NMR (CDCl$_3$) δ 9.02 (d, J = 2.1 Hz, 1H, H$_3$ of DNP), 8.46 (s, 1H, H$_6$), 8.07 (dd, J = 9, 2.4 Hz, 1H, H$_5$ of DNP), 7.35-7.09 (m, 10H, H6 of DNP and Ar), 6.75-6.64 (m, 4H, Ar),
35   6.41 (dd, J = 9, 2.4 Hz, 1H, H$_{1'}$), 4.70 (m, 1H, H$_{3'}$), 4.16 (m, 1H, H$_4$), 3.73 (2s, 6H, 2CH$_3$), 3.38 (d, J = 7.5 Hz, 2H, H$_{5'}$), 2.64 (m, 1H, H$_{2'}$), 2.44 (m, 1H, H$_{2'}$); UV (0.1 M TEAA, pH 7.1 with 82% CH$_3$CN)

SUBSTITUTE SHEET (RULE 26)

264, 324 nm; HR FAB MS calculated for $C_{36}H_{32}N_4O_{11}S$ $(M)^+728.1788$ found 728.1802 (1.9 ppm). (Synthesized by the method of Bleasdale et al. (Bleasdale *et al.*, 1990)).

Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-S-(2,4-dinitrophenyl) mercapto-2'-deoxyuridine-3'-O-(2-cyanoethyl-N,N'-diisopropyl) phosphoramidite (Compound 7).

Compound 6 (13 mg, 18 μmol) was dried under vacuum in the reaction vessel. Anhydrous CH_3CN (0.3 mL) was added and the reaction vessel was transferred to an atmosphere bag filled with dry N_2. A mixture of 2-cyanoethyl-N,N,N'N'-tetraisopropyl-phosphorodiamidite (5.5 μL, 17 μmol) and tetrazole (1.3 mg, 18 μmol) in anhydrous CH_3CN (50 μL) was added. The reaction mixture was maintained at room temperature for 30 min. An additional mixture of 2-cyanoethyl-N,N,N'N'-tetraisopropylphosphorodiamidite (2.3 μL, 7.2 μmol) and tetrazole (0.6 mg, 8.3 μmol) in anhydrous CH_3CN (25 μL) was added and the reaction mixture was stirred for 30 min. CH_2Cl_2 (20 mL) and 5% (w/v) aq. $NaHCO_3$ (15 mL) were added to the mixture; the organic layer was removed and re-extracted with water, dried (Na_2SO_4) and evaporated. The residue was subjected to column chromatographic purification [CH_2Cl_2:MeOH:pyridine(99.5:0:0.5 to 98.5:1.0:0.5)] to obtain Compound 7 (8.6 mg, 52%). The yellow product was analyzed by analytical RP HPLC using a gradient of buffer A and acetonitrile [70% (v/v) CH_3CN to 80% in 3 min.; 80% CH_3CN to 100% in 15 min.]. UV absorption was monitored at wavelengths of 254 and 330 nm. The diastereomers eluted at 11.09 and 12.49 min. UV scans of the diastereomers showed absorption peaks at 264 and 327 nm. Phosphorus NMR indicated the presence of two diastereomers [^{31}P NMR $(CDCl_3)$ δ 149.12, 149.27 ppm].

Example 2. Automated Synthesis of an Oligonucleotide Containing 5-SDNP-dU

Synthesis and Purification of Oligonucleotides.

Oligonucleotides were synthesized on an Applied Biosystems 392 DNA Synthesizer using the standard β-cyanoethyl-protected phosphoramidite method. A series of three modified oligonucleo-

tides were prepared on a 40 nmole scale using a Trityl On
method. The thiol-protected analog, Compound 7, was manually
dissolved into anhydrous acetonitrile at a concentration of
0.027 M and the solution was attached to the spare port on the
5   DNA synthesizer. As the concentration of Compound 7 was lower
than the concentration of the commercial phosphoramidites used
in these syntheses (0.05 M), the coupling time for the modified
analog was increased by 120 seconds. Syntheses of the
corresponding unmodified oligonucleo-tides were also performed.
10   The sequences of the oligonucleotides are shown in Table 1.

The oligonucleotides were removed from the column by
treatment with concentrated NH$_4$OH(1 mL). Deprotection of the
exocyclic amines was then accomplished by treatment of the
oligonucleotides with an additional 1 mL of fresh concentrated
15   NH$_4$OH at room temperature for 44 hours. The NH$_4$OH solution was
concentrated using a Savant Speed Vac without heat. Small
portions of Et$_3$N were periodically added to the concentrating
solution to maintain a basic pH. The crude oligonucleotides
were re-dissolved in 10 mM TEAB (pH 7.1) and purified as the DMT
20   ethers using semi-preparative RP HPLC. Two different gradients
of buffer B and acetonitrile were utilized. The shorter
oligonucleotides (7 and 22 nucleotides) were purified with an
elution gradient of CH$_3$CN from 5% (v/v) to 30% in 15 min. then
30% to 100% in 5 min. The 24-mer oligonucleotides were purified
25   with an elution gradient of CH$_3$CN from 5% to 30% in 18 min. then
30% to 100% in 5 min. The appropriate fractions were evaporated
to dryness using a Jouan Concentrator without heat. Following
purification of the oligonucleotides, the DMT group was removed
by treatment with 80% (v/v) HOAc (200 μL) at room temperature
30   for 20 min. The oligonucleotides were ethanol precipitated and
re-dissolved into 300 μL of nanopure water and stored at -20°C.
The purity of the oligonucleotides was analyzed by analytical RP
HPLC (Figure 2A, Oligonucleotides A and B shown). Elution of
the oligonucleotides was performed using a linear gradient of
35   CH$_3$CN (5 to 20% in 15 min.) with buffer B. The purified
oligonucleotides were also analyzed by 25% polyacrylamide-7 M
urea gel (acrylamide/methylene bisacrylamide = 19/1) (Figure 3A,
Oligonucleotides C-F shown).

Enzymatic Digestion with Nuclease P1 and CIP. The oligonucleotides A and B were subjected to enzymatic digestion. The pure oligonucleotides were incubated with Nuclease P1 (2 U) at 37 °C for 6 hours in a total volume of 50 μL of buffer C.

5   Five microliters of 10X buffer D and 43 μL of H_2O were added to the reaction. CIP (2 U) was added and the pH of the reaction was adjusted with 5% (v/v) NaOH (0.8 μL). The reaction mixture was incubated overnight at 37°C. To recover the nucleosides, 11 μL of 3 M NaOAC, pH 5.3 and 250 μL of absolute EtOH were added

10   to the reaction mixture. The mixture was stored at -80°C for 4 hours and then centrifuged at 14000 rpm for 20 min. The supernatant fraction was removed and concentrated. The resulting residue was re-dissolved into 35 μL of buffer A and analyzed by analytical RP HPLC (Figure 2B). The elution

15   gradient consisted of buffer A with an increasing concentration of CH_3CN (0% to 5% in 5 min., 5% to 6% in 15 min., 6% to 50% in 5 min. and finally 50% to 80% in 15 min.). The UV absorption was monitored at 260 and 330 nm. Commercial nucleosides and monophosphates of dA, dC, dG and T were utilized as standards.

20   Synthetic 5-S-(2,4-dinitrophenyl) mercaptodeoxyuridine (Compound 5) was also utilized as a standard. An enzymatic digestion without oligonucleotide was performed to identify background peaks arising from the buffer and enzymes.

Deprotection and alkylation of a DNP-modified oligonucleotide.

25   Oligonucleotide F (100 pmol) was radiolabeled by reaction with T4 Polynucleotide Kinase using [γ-^{32}P]ATP. The labeled oligonucleo-tide F was isolated by ethanol precipitation and re-dissolved into 30 μL H_2O. The oligonucleotide (4.5 μL) was treated with 1.4 M BME either at 45 °C for 4 hours or at room

30   temperature for 12 hours in a 20 μL solution of 20 mM Tris-HCl, pH 8.2. Following this deprotection to remove the DNP group, half of the reaction mixture was precipitated with ethanol while the other half of the mixture was subjected to alkylation conditions with 5-IAF. The corresponding unmodified

35   oligonucleotide E was treated identically to serve as a control.
        To recover the deprotected oligonucleotides by ethanol precipitation, 0.36M NaOAc, pH 5.6 (100 μL), 1M $MgCl_2$(1.2 μL) and

nanopure water (10 μL) were added to the reaction mixture (10 μL) to provide a final volume of 121 μL. Absolute ethanol (300 μL) was added and the mixture was cooled at -80°C. The oligonucleotides were pelleted by centrifuging at 14000 rpm for

5   20 min. The supernatant was removed and the pellet was washed with cold 95% (v/v) ethanol. The mixture was chilled for 20 min. in a dry ice/ethanol bath and spun at 14000 rpm for 20 min. Following the removal of the supernatant fraction, the recovered oligonucleotides were dissolved into 20 μL of nanopure water.

10   Mixtures of oligo-nucleotides (4.5 μL), 100 mM DTT (0.5 μL) and buffer E (5 μL) were prepared and analyzed by electrophoresis on a 25% polyacrylamide-7 M urea gel (acrylamide/methylene bisacrylamide = 19/1) (Figure 3B).

       Alkylation of the deprotected thiol moiety was accomplished

15   by the addition of 15 μL of 20 mM Tris-HCl (pH 4.8) to 10 μL of the deprotection reaction (pH 8.2). Following this addition, the pH of the alkylation reaction was 7 (as judged by pH paper). One microliter of 10mM 5-IAF in DMF was added and the reaction mixture was maintained at room temperature overnight. The

20   oligonucleotides were recovered by ethanol precipitation and re-dissolved into 20 μL nanopure water. The alkylated oligonucleotides (4.5 μL) were mixed with 100 mM DTT (0.5 μL) and buffer E (5 μL) and analyzed by electrophoresis on a 25% polyacrylamide-7M urea gel (acrylamide/methylene bisacrylamide

25   = 19/1) (Figure 3B).

       DNA Synthesis via PCR using a DNP-modified oligonucleotide.
       PCR reactions were performed on a Perkin Elmer Cetus Gene Amp PCR System 9600. The DNA fragment containing the bacteriophage lambda PR' promoter, 6S gene, and tR' terminator,

30   with the lambda gut site (Yang et al., 1987) was amplified from plasmid pHA100 (Zhang & Hanna, 1995) using either oligonucleotide E or F and the universal T7 primer (Stratagene). The DNA product was purified by excision of the band from an agarose gel and extraction of the DNA using the commercially

35   available GENECLEAN kit (BIO 101 Inc., La Jolla, CA). The PCR products were dissolved into equal volumes of buffer F and analyzed by electrophoresis on a 1.5% agarose gel. Detection of

DNA was achieved by addition of ethidium bromide to the electrophoresis buffer. (Figure 4A).

Immunodetection of DNP-labeled DNA.

Dot blot experiments were performed manually. The DNP-labeled
5 PCR product and the corresponding unlabeled DNA fragment (1 pmole each) were in 10X SSC solution (1.5 M NaCl, 0.15 M Na₃Citrate, pH7). Samples (13 µL) were applied to a pre-wetted charged nylon membrane (Zeta Probe GT; Biorad). The membranes were air dried, rinsed briefly with 2X SSC and heated at 80°C
10 under vacuum for 30 min. The membranes were blocked by overnight incubation at 4°C in TBS-T (20 mM Tris HCl, 500 mM NaCl, 0.4% (v/v) Tween 20, pH 7.6) containing 10% (w/v) dried milk. The membranes were then washed with TBS-T two times briefly, one time for 15 min., and two times for 5 min. The
15 membranes were incubated with a 1:1500 dilution of anti-DNP antibodies for 1 hour, washed as described above and then incubated with a 1:10,000 dilution of horseradish peroxidase-linked whole anti-rabbit Ig for 1 hour. Membranes were then washed one time for 15 min. and 4 times for 5 min. with TBS-T.
20 Detection of the antibodies was carried out using the ELC chemiluminescence method (Amersham). The membranes were incubated with the detection agents for 1 min. before exposure to x-ray film for 2 min. (Figure 4B).

RESULTS

25   The synthesis and characterization of a novel deoxyuridine phosphoramidite (Compound 7) containing a protected thiol group at the 5 position of the mercaptopyrimidine ring is described (Figure 1). This DNP-labeled analog was site-specifically incorporated internally into a series of oligonucleotides.
30 Significant differences in overall synthetic yields were not observed between oligonucleotides made with normal deoxynucleoside phosphoramidites and those which contained the analog, indicating that the modified phosphoramidite is a suitable reagent for automated oligonucleotide synthesis.
35 Enzymatic digestion of the DNP-labeled oligonucleotide established the stability of the modified pyrimidine to the

conditions for chemical synthesis and purification of oligonucleotides (Figure 2B). Selective deprotection of the 5-thiol moiety with BME and subsequent modification of the unmasked thiol with 5-IAF were verified by gel electrophoresis of the oligonucleotides (Figure 3B). Incorporation of the DNP-labeled oligonucleotide into double-stranded DNA was achieved using PCR. Isolation and characterization of the PCR products included gel electrophoresis and immunodetection with anti-DNP antibodies (Figure 4A and B).

Synthesis.

Synthesis of the deoxyuridine phosphoramidite analog (Compound 7) was carried out following the sequence of reactions outlined in Figure 1. Compounds 2-7 have not been previously described in the literature. The method chosen for incorporation of the thiol group required thiocyanation of the pyrimidine ring. Initially, this reaction was performed on deoxyuridine and the reaction conditions utilized were similar to those previously described (Bradley & Hanna, 1992; Nagamachi et al., 1974; Torrence et al., 1968). Although literature reports indicate yields of 54% for the thiocyanation of deoxyuridine, this was difficult to achieve. The yield for this reaction varied considerably and was at times as low as 9%. A major side product in this reaction was isolated and identified by NMR and IR as 5'-O-acetyl-5-thiocyanatodeoxyuridine indicating that along with thiocyanation, acetylation of the sugar ring had occurred.

To circumvent the problems associated with the reaction described above, 3',5'-O-diacetyl-2'-deoxyuridine (Compound 1) was utilized as the starting material. Protection of the sugar hydroxyls by acetylation prior to thiocyanation results in higher yields (95%) for the reaction (Nagamachi et al., 1974). However, removal of the acetyl protecting groups could not be accomplished using standard basic conditions required for deesterification without destruction of the thiocyanato group. Therefore, a synthetic approach was developed which required reduction of the thiocyanato compound to the mercapto analog prior to removal of the acetyl groups (Figure 1). Thiocyanation

of the 3',5'-O-diacetyl-2'-deoxyuridine (1) was achieved following the literature procedure described above except that the crude material was extracted with petroleum ether to remove nonpolar impurities prior to column chromatographic purification. The pure product was obtained in 82% yield as a white solid.

Reduction of the thiocyanato derivative 2 to the mercapto analog (Compound 3) was achieved by reaction with DTT in a solvent mixture of MeOH and 0.1 M EDTA (pH=7.8). This reaction is a variation of a reported procedure (Lin et al., 1988) for the reduction of 3'-azido-2',3'-dideoxy-5-thiocyanatouridine; this mercapto analog however was not isolated. Purification of Compound 3 was achieved by acidification of the reaction mixture followed by concentration of the mixture. The residue was extracted with CH_2Cl_2 and H_2O. The presence of the acetyl groups insured the product was organic solvent soluble. In order to remove residual DTT, the organic layer was reduced and the residue triturated with water to crystallize the product as well as to dissolve residual DTT. Compound 3 was obtained as a white solid in 75% yield.

The mechanism for the reduction of the thiocyanato requires displacement of the cyano group by DTT. Since no attempt was made to remove the cyanide ion before acidification of this reaction mixture, every step of the isolation of Compound 3 from acidification to trituration with H_2O was performed in a fume hood. This synthesis of Compound 3 was designed for the preparation of both 5-alkyldithio and 5-alkylthio or 5-arylthiopyrimidine analogs. To prevent degradation of the 5-alkyldithio compounds by DTT, purification of Compound 3 was necessary. This is not the case with formation of 5-alkylthio pyrimidines (Lin et al., 1988) and this purification step has been eliminated for Compound 4 resulting in a slight decrease in yield.

The DNP protecting group has been employed in oligonucleotide synthesis as a protecting group for 6-mercaptopurine and 6-thioguanine phosphoramidites (Xu et al., 1992a; Xu et al., 1992b) and as a functional moiety on a non-nucleoside phosphoramidite (Grzybowski et al., 1993). These

studies indicated the DNP group was stable during automated DNA
chemical synthesis. Conversion of mercaptan 3 into the 2,4-DNP
analog 4 was achieved by reaction with 2,4-dinitrofluorobenzene
in anhydrous CH_3CN with Et_3N present as a catalyst (Xu et al.,

5   1992b). The product was isolated as a yellow solid in 88%
yield. Attachment of the thiol protecting group prior to
removal of the acetyl moieties insured that the product was
soluble in organic solvents, which aided in the purification of
the material. More importantly, protection of the sulfur

10   inhibits the oxidation of the analog to the disulfide compound,
a potential reaction of 5-mercaptopyrimidines (Kalman & Bardos,
1967).

       Another method of preparing 5-arylthiol ethers of
deoxyuridine has been previously described. Bergstrom et al.

15   has prepared a series of alkyl and aryl mercaptodeoxyuridine
analogs using a palladium-mediated reaction between 5-
(chloromercuri)-2'-deoxyuridine and the appropriate disulfide
(Bergstrom et al., 1991). Although this method offers a simple
route to arylthiol ether analogs it was not selected for study.

20   Development of a more versatile intermediate such as Compound 3
which can be easily converted into various types of analogs
including disulfides was desired for future studies.
Additionally, the synthetic utility of the method has not been
fully evaluated and therefore formation of the desired

25   dinitrophenyl analog (Compound 5) was uncertain.

       Deesterification of nucleosides is commonly performed under
basic conditions using NH_3 or NaOMe (Kini et al., 1989; Trivedi
& Bruns, 1989). Literature reports suggest the DNP group is
marginally stable to nucleophiles such as NaOMe; the stability

30   of this group being strongly dependent on the reaction
conditions (Greene, 1981). Successful removal of the acetyl
groups was achieved by reaction of the diacetyl analog (Compound
4) with NaOMe in MeOH without an apparent loss or modification
of the DNP group as judged by NMR and IR. Following the

35   deacetylation, the diol (Compound 5) was isolated as a yellow
solid in 81% yield. An initial attempt to deacetylate Compound
4 under mild reductive methods was performed (Brown et al.,
1982; Soai & Ookawa, 1986). However, reaction with $LiBH_4$ in a

mixture of Et$_2$O and MeOH did not result in the formation of Compound 5, possibly due to the poor solubility of the starting material.

Incorporation of the DMT protecting group at the 5' position of nucleosides is typically accomplished by reaction with DMTCl in the presence of pyridine or dimethylpyridine (Chaudhary & Hernandez, 1979). However, as utilization of these conditions with other 5-thiol modified nucleosides was not successful (unpublished results), this reaction was not attempted for the synthesis of Compound 6. Conversion to the 5'-dimethoxytrityl analog (Compound 6) was achieved by reaction of Compound 5 with DMT$^+$BF$_4$- in the presence of DBMP in CH$_3$CN at reflux. This reagent, DMT$^+$BF$_4$, has only recently been utilized in the preparation of 5' protected nucleosides (Bleasdale et al., 1990; Lakshman et al., 1992).

The proposed mechanism for dimethoxytritylation with DMTCl and pyridine requires formation of a dimethoxytrityl pyridinium salt which then undergoes reaction with the 5' hydroxyl group (Chaudhary & Hernandez, 1979). Formation of the pyridinium salt is considered the limiting step. The reagent DMT$^+$BF$_4$ serves as a source for the dimethoxytrityl cation. By preforming the cation prior to reaction, the slow step in the dimethoxytritylation has been eliminated and the ether should be more easily formed. Although Compound 6 was obtained using this reagent, the reaction was performed under reflux conditions. It is unclear why the reaction requires such stringent conditions but presumably steric hindrance plays some part. Modification of other bases has been reported to inhibit dimethoxytritylation possibly due to steric hindrance (Cosstick & Douglas, 1991).

Silica gel column chromatography was employed to isolate Compound 6. Due to the acidic nature of the silica gel, a small percentage of pyridine was added to the eluting solvent to prevent loss of the acid sensitive DMT group. Initially, Et$_3$N was used in the solvent system; however triethylammonium salts of Compound 6 were isolated as determined by NMR. This was not observed during the synthesis of the standard 5'-dimethoxytrityl-2'-deoxyuridine. Switching to pyridine, a weaker base, in the eluting solvent system eliminated this

SUBSTITUTE SHEET (RULE 26)

problem. A satisfactory high resolution FAB mass spectra was obtained for Compound 6.

The Compound 6 was isolated as a yellow solid in yields ranging from 35 to 48%, making formation of the dimethoxytrityl ether the lowest yielding reaction in this synthetic sequence. However, approximately 20% of the starting material, Compound 5, was converted into the 3',5'-bis(dimethoxytrityl) analog, which was isolated and converted back into 4 for recycling into the synthesis.

The final reaction in the preparation of Compound 7 is the incorporation of the phosphoramidite moiety at the 3' position of the nucleoside. This is commonly accomplished by reaction of the appropriately protected nucleoside with 2-cyanoethyl-N,N-diisopropyl chlorophosphoramidite in the presence of the N,N-diisopropylethylamine (DIEPA) (Sinha et al., 1984). Utilization of these conditions did not result in the desired product and starting material was re-isolated. However, a standard reaction with 5'-dimethoxytrityldeoxyuridine was successful. A second method for incorporation of the phosphoramidite moiety which utilizes   the   reagent   2-cyanoethyl-N,N,N,'N'-tetraisopropylphosphorodiamidite in the presence of tetrazole in CH$_3$CN was evaluated (Barone et al., 1984). Unlike the basic reaction with DIEPA, these slightly acidic conditions provided the DNP-labeled phosphoramidite, Compound 7. RP HPLC analysis of Compound 7 indicated the presence of a pair of closely eluting peaks (11.0 and 12.4 min) representative of the expected 1:1 mixture of diastereomers. The overall purity of the sample was determined to be greater than 96% and no starting material was observed in the chromatogram. HPLC data was quantified using a simple area percent method. Similar results were observed using silica gel TLC analysis (CH$_2$Cl$_2$:MeOH:Et$_3$N;96:3:1). The phosphorous NMR spectrum showed two characteristic peaks associated with the diastereomeric phosphoramidites (149.12 and 149.27 ppm). An additional peak appeared in the NMR at 14.3 ppm. Based on the chemical shift of this peak, the impurity is presumed to be a hydrolysis product (Scremin et al., 1994). No attempt was made to remove this material prior to oligonucleotide synthesis as it is not reactive in the

36

oligonucleotide chemistry. A high resolution FAB mass spectrum
could not be obtained for Compound 7 due to its instability
during analysis.   This was not surprising as the standard
deoxyuridine phosphoramidite displayed a high degree of
5   instability during high resolution FAB MS analysis; however, a
small molecular ion peak was identified in the spectrum.
Evaluation of Compound 7 using low resolution mass spectroscopy
showed a fragment (m/z=407) consistent with the loss of the DMT
and the phosphoramidite moieties. As a similar fragmentation
10   pattern was observed for the standard deoxyuridine
phosphoramidite under these experimental conditions, the
information supports the proposed structure.

Syntheses and purification of oligonucleotides.
Internal incorporation of the phosphoramidite analog,
15   Compound 7, into series of oligonucleotides has been
accomplished using an automated DNA synthesizer. The natural
oligonucleotides with corresponding sequences were prepared as
standards.   The sequences for the oligonucleotides which are
listed in Table 1 were chosen for two reasons: the 7-mer was
20   designed for characterization purposes while the 22-mer and the
24-mer were developed to study the DNA-protein interactions
during transcription of the bacteriophage lambda 6S gene. The
concentration of the phosphoramidite solution of Compound 7 was
diluted twofold as compared to the normal phosphoramidites (dA,
25   dG, T and dC) utilized in these DNA syntheses. This dilution
was performed to conserve material. The coupling time for the
reaction involving the modified phosphoramidite was increased by
120 seconds.   Under these reaction conditions, the coupling
yield for the modified phosphoramidite was >96% which was
30   comparable to the coupling yields of the other standard
phosphoramidites.   The average stepwise yields for all six
oligonucleotide syntheses were determined automatically by the
ABI synthesizer using a trityl cation assay. These yields are
listed in Table 1.
35   Selective removal of the exocyclic amine protecting groups
in the presence of the DNP moiety was achieved by treatment with
NH_4OH at room temperature for 44 hours. Typically, this

SUBSTITUTE SHEET (RULE 26)

deprotection is performed at elevated temperatures of 55°C for
8 hours (ABI technical bulletin #13, 1987). However, studies
involving reaction of the DNP-labeled nucleoside 5 with conc
NH_4OH at 55°C indicated the analog was not stable to these

5   conditions but was stable at room temperature (data not shown).
As successful removal of standard exocyclic amine protecting
groups at room temperature has been reported for other
oligonucleotides containing modified phosphoramidites, these
conditions were utilized (Sheardy & Seeman, 1986). Although

10   these initial studies were performed with the standard exocyclic
amine protecting groups, phosphoramidites containing a wide
variety of protecting groups which can be efficiently removed
under mild conditions are now commercially available. These
will be investigated for compatibility with phosphoramidite

15   Compound 7.

The crude oligonucleotides were purified as the DMT ethers
by RP HPLC. Removal of the DMT group under standard acidic
conditions did not affect the DNP moiety. Following removal of
the DMT group, the oligonucleotides were analyzed for purity by

20   both analytical RP HPLC and gel electrophoresis as shown in
Figures 2A and 3A respectively. The oligonucleotides showed one
major product in both analyses. The retention times of the DNP-
labeled oligonucleotides were greater than the corresponding
natural oligonucleotides in the HPLC analyses due to the

25   lipophilic nature of the dinitrophenyl group. The RP HPLC
retention times for the purified oligonucleotides are listed in
Table 1. The overall yields of the pure material which were
calculated by dividing the optical density of pure
oligonucleotide by the optical density of crude DMT-labeled

30   oligonucleotides times 100% are found in Table 1 (Kuimelis &
Nambiar, 1994). The yields for the corresponding normal and
DNP-labeled oligonucleotides were comparable.

Enzymatic digestion of the heptanucleotide.

To establish that the DNP group is stable to the conditions

35   of automated DNA synthesis, deprotection, and purification, a
DNP-labeled oligonucleotide was enzymatically digested into its
corresponding nucleosides and evaluated by RP HPLC. In order to

SUBSTITUTE SHEET (RULE 26)

observe the single DNP-labeled nucleoside present in the enzymatic digest, the short chain oligonucleotide B was used for this analysis. The sequence chosen for this heptanucleotide did not contain deoxycytidine. Deoxycytidine co-elutes with 5-
5   mercapto-deoxyuridine, a potential side product arising from the loss of the DNP group, in the gradient system utilized in the analysis. Deoxycytidine was therefore omitted from the sequence to eliminate uncertainty in identifying the nucleoside peaks. 5-Mercaptodeoxyuridine was synthetically prepared in this lab by
10   reduction of 5-thiocyanatodeoxyuridine (Nagamachi et al., 1974).

Initial degradation of the oligonucleotide B and the corresponding normal oligonucleotide A was attempted by concurrent digestion with snake venom phosphodiesterase (SVP) and CIP (Kuimelis & Nambiar, 1994). This method proved useful
15   in the digestion of the natural oligonucleotide but did not appear to fully degrade the DNP-labeled material. This is consistent with other studies in which SVP has failed to function with some modified oligonucleotides and nucleotides (Gao et al., 1992; Hanna et al, 1989; Hanna et al., 1993).
20   Substitution of the exonuclease SVP with an endonuclease Nuclease P1 resulted in complete digestion. Due to a difference in the pH requirements of these enzymes for activity, this digestion was accomplished in two consecutive enzymatic reactions. First, Nuclease P1 cleaved the phosphodiester bonds
25   to provide the 5'-monophosphates which were then subjected to treatment with CIP. The resulting nucleosides were isolated and analyzed by RP HPLC. The results are shown in Figure 2B.

This enzymatic degradation provided the correct ratios of unmodified and modified nucleotides. For oligonucleotide A, the
30   ratio of dA, dG and T was 1.9:2.1:2.9. For oligonucleotide B, the nucleoside 5 is also expected in the reaction mixture providing a ratio of dA, dG, T and 5 of 1.9:2.1:1.9:1.1. Chemically synthesized Compound 5 was utilized as a standard. Both the enzymatically obtained and the chemically obtained
35   nucleoside Compound 5 eluted with the same retention time during RP HPLC analysis. In addition, these compounds provided identical UV spectra with absorption peaks appearing at 262 and 334 nm. Finally, no unidentifiable peaks were observed in the

SUBSTITUTE SHEET (RULE 26)

HPLC chromatograms of the enzymatic digests and only one peak representing nucleoside Compound 5 appeared in the HPLC chromatogram monitored at 330 nm. In this analysis, absorption at 330 nm is associated with the presence of either the dinitro-
5   phenylthiol ether or the 5-mercaptopyrimidine moiety. If the DNP-labeled analog had undergone modification during automated DNA synthesis, deprotection, or purification, extraneous peaks were predicted to appear in this chromatogram.

Deprotection and alkylation of a DNP-modified oligonucleotide.
10   Removal of the DNP moiety was accomplished by reaction of oligonucleotide F with 1.4 M BME in Tris buffer pH 8 (Shaltiel, 1967). The reaction was performed under two different reaction conditions: room temperature for 12 hours and 45°C for 4 hours. Oligonucleotide E was reacted under the same conditions to serve
15   as a standard. No attempt was made to isolate the deprotected oligonucleotide prior to modification with 5-IAF. The pKa of BME is 9.6 while the pKa of 5-mercaptodeoxyuridine is 5.0 (Bardos & Kalman, 1966). Modification of a thiol by reaction with haloacetamido compounds such as 5-IAF requires formation of
20   the thiolate anion. By lowering the pH of the reaction mixture, the selectivity for modification of the oligonucleotide is enhanced over that of BME, a potential side reaction. The oligonucleotide products in both the deprotection and modification reactions were isolated by ethanol precipitation
25   and analyzed by electrophoresis using a 25% polyacrylamide-7 M urea gel. The results are shown in Figure 3B. Lanes 1 through 4 represent reactions utilizing oligonucleotide E. No difference is observed in these four lanes indicating the natural oligonucleotide structure is stable to the conditions
30   needed for deprotection and alkylation of the thiol. Lanes 5 and 7 represent reactions of oligonucleotide F with BME at room temperature and 45°C respectively, while lanes 6 and 8 depict the corresponding alkylation reactions which were carried out at room temperature. No apparent mobility shift is observed
35   between the natural oligonucleotide (lanes 1-4) and lanes 5 and 7, which is consistent with our observations that a 20-mer RNA containing 5-SH-UTP co-migrates in this gel system with 20-mer

40

containing no analog (He et al., 1995). In contrast, a substantial difference is observed following modification with fluorescein, consistent with other reports of gel mobility shifts of fluorescein-labeled oligonucleotides (Clegg et al.,

5   1992; Mergny et al., 1994). In fact, the selection of 5-IAF for the alkylation experiments instead of a thiol-modifying photocrosslinking reagent such as p-azidophenacyl bromide (APB) was due in part to the ease in which the reaction results could be monitored by gel electrophoresis. The conditions for

10   alkylation of the oligonucleotide with 5-IAF however should prove useful with other haloacetyl derivatives such as APB. We have specifically labeled the thiol group in an RNA molecule containing 5-SH-UMP (He et al., 1995) with APB. We have also attached APB to a DNA molecule generated by PCR with a

15   oligonucleotide containing (compound 7), after BME removal of the DNP group (not shown).

      As the mobility shift associated with the DNP label is small (Figure 3A), it is possible that a mixture of DNP-labeled and 5-mercapto-labeled oligonucleotides may appear as one band

20   on a 25% polyacrylamide-7 M urea gel. Therefore, one can not determine the extend to which deprotection has occurred. However, the appearance of a new band in lanes 6 and 8 (Figure 3B) with the subsequent disappearance of the band seen in lanes 5 and 7 indicates that the oligonucleotide underwent

25   deprotection and alkylation. Although deprotection does occur with time at room temperature, the reduced time required to deprotect at 45°C makes this the preferred method.

PCR reaction using a DNP-modified oligonucleotide.

      The DNP-labeled oligonucleotide was incorporated into

30   double stranded DNA using PCR. Plasmid pHA100 served as the template for this reaction. A 492 base pair segment of this plasmid was amplified using the universal T7 primer and either oligonucleotide E or F. The products isolated from the reaction were analyzed on a 10% polyacrylamide-7 M urea gel (data not

35   shown) and a 1.5% agarose gel (Figure 4A). Differences in mobility were not observed between the standard PCR product and the DNP-labeled PCR product in these gel analyses and no

41

difference in yield was observed.

Immunodetection of DNP-labeled oligonucleotides and DNA.

To insure the stability of the DNP protecting group during
the PCR reaction, the products described above were subjected to
5   immunodetection studies using rabbit polyclonal anti-DNP
antibodies.   The natural PCR product served as a negative
control.   The DNP group was detected using a chemiluminescence
method and only the DNP-labeled nucleic acid provided a signal
(Figure 4B).

10   Example 3: Analysis of Protein-DNA Interactions
                        Utilizing 5-S-DNP-dU

Photochemical crosslinking is a powerful technique for
characterization of   both   RNA-protein   and   DNA-protein
interactions   in   nucleoprotein   complexes.   Photochemical
15   crosslinking "traps" weak or transient associations which might
not survive isolation procedures such as immunoprecipitation,
gel filtration, or filter binding.   There are a number of
approaches which can be taken in photochemical crosslinking.
One involves direct irradiation of a nucleoprotein complex with
20   short wavelength ultraviolet light.   This method relies upon the
direct excitation of nucleotides or amino acids to generate
chemically reactive species.   Such nonspecific labeling can
reveal whether a protein-nucleic acid interaction exists, but
there are several problems associated with this approach. These
25   include low crosslinking yields, the inability to incorporate a
crosslinking group at a specific site in the nucleic acid, and
degradation of some proteins during irradiation.   A second
approach involves the use of nucleotide analogs modified with
photoreactive crosslinking groups (reviewed in M. Hanna, 1989,
30   Methods in Enzymology).   These groups are chemically inert in
the absence of light but are converted to chemically reactive
species upon irradiation.   Often this can be achieved with long
wavelength ultraviolet or visible light, resulting in less
nicking of protein or nucleic acid.   Molecular contacts can be
35   identified at the level of specific nucleotides in nucleic acids
and amino acids in proteins.   One can therefore gain structural

SUBSTITUTE SHEET (RULE 26)

information about macromolecular complexes which cannot be obtained by most other biochemical approaches.

For such mechanistic studies a variety of nucleotide analogs have been developed and utilized (M. Hanna 1989, Methods in Enzymology; Bradley and Hanna, 1992). These analogs are either tagged with a photoreactive crosslinking group or contain functional groups that can be tagged with crosslinkers after incorporation into DNA or RNA. These analogs can be incorporated into nucleic acids, either enzymatically or chemically, to analyze molecular interactions in protein-nucleic acid complexes. Upon photoactivation, analog-tagged nuecleic acids become covalently attached to adjacent macromolecules (protein, DNA, RNA) with which they have direct interactions. Therefore, nucleic acid binding domains in complexes can be characterized, nucleic acid binding proteins in extracts can be identified, and determinants involved in these specific interactions can be characterized at the level of individual nucleotides and amino acids. This methodology allows mechanistic questions about the way that specific protein-DNA or protein-RNA interactions regulate gene expression to be asked. Herein described are the approaches used to characterize the protein-DNA interactions in the E. coli transcription complex which forms on the bacteriophage lambda $P_{R'}$ promoter.

Preparation of Site-Specifically Modified DNA Oligonucleotides
The method described here is used to probe for interactions between a specific nucleotide in DNA and a DNA-binding protein. In this case, interactions between a specific nucleotide on the non-template strand of the $P_{R'}$ promoter are sought. The chemical approach to preparation of site-specifically modified DNA involves incorporation of a modified nucleotide during the automated synthesis of an oligonucleotide. Because the conditions used in automated oligonucleotide synthesis are rather harsh, the direct incorporation of a photocrosslinking azide group by this method is not possible, and a means for the post-synthetic incorporation of the azide group is required. 5-S-DNP-dU was incorporated into a DNA oligonucleotide, as described herein. For analysis of protein interactions with the

-12 position in the nontemplate strand of the p_R, promoter, an oligonucleotide with the sequence 5'-5D-DNP-dU-AA ATT TGA CTC AAC GAT GGG (SEQ. ID. NO:1) was synthesized on an Applied Biosystems 392 DNA Synthesizer using the standard β-cyanoethyl-protected phosphoramidite method by attaching the analog to the fifth substrate port. After synthesis, removal of the oligonucleotide from the column, and removal of the exocyclic amine protecting groups, the oligonucleotide was ethanol precipitated, dissolved in water and stored at -20°C.

Analysis of 5-S-DNP-dU Effects on Hybridization Properties of ODNs

    To evaluate the effect of incorporation of one or more 5-SDNP-dU analogs into oligodeoxynucleotides (ODNs) on their hybridization properties, the ODNs shown in Table 2 were synthesized. The observed melting temperatures (T_m) showed that incorporation of one DNP analog decreased the melting temperature by only 1.8°C (hybrid V), compared to the normal DNA hybrid (I), and substitution with two DNP analogs caused a decrease of only 4.4°C (hybrid VI). Therefore, the effect of a single substitution is less than that caused by a single mismatch involving unmodified nucleotides (hybrids II, III and IV vs I), and the effect of two modified analogs is still less than that of a single T-C mismatch (hybrid IV). Mismatches of 2 or more are commonly used for site-directed mutagenesis, which involves hybridization of the oligonucleotides to single-stranded DNA. Substitution with the DNP analog should therefore work as well, if not better. The use of ODNs containing one or more 5-SDNP-dU analog in assays requiring specific hybridization to a complementary strand is therefore quite feasible.

    By comparison, oligodeoxynucleotides (30-mers) containing a previously described nucleobase fluorescent derivative, 5-Amido-(Carboxyfluorescein)-2'-deoxyuridine, caused decreases in melting temperature of 2°C for ODNs with one nucleobase substitution, 6°C for ODNs with two nucleobase substitutions, and 12°C for ODNs containing three nucleobase substitutions (Jadhav et al. 1997). A more serious effect on the hybridization properties was observed for another previously

described analog, 2-thiodeoxyuracil. Placement of even a single 2-thiodeoxyuracil analog into oligodeoxynucleotides (14-mers) caused a decrease in T_m of 3-4°C (Kuimelis and Nambiar, 1994).

Table 2. T_m values for native and modified ODNs

ODN 1 5' (SEQ. ID. NO:2)	GAC	VCA	ACG	AWG	GGX	TAA	TYC	G 3'
ODN 2 3' (SEQ. ID. NO:3)	CTG	AGT	TAC	TZC	CCA	ATT	AAG	C 5'
ODN	T	1	1	1	2		
Hybrid	V	W	X	Y	Z	T_m(°C)	Drop in T_m^a
I	T	T	T	T	A	67.3	0
II	T	1	T	1	G	64.5	2.6
III	T	T	T	T	T	63.6	3.5
IV	T	T	T	T	C	61.5	5.6
V	T	5^b	T	T	A	65.5	1.8
VI	T	5	5	T	A	62.9	4.4
VII	T	5	T	5	A	62.5	4.8
VIII	5	5	T	5	A	58.0	9.3

a - as compared to unmodified duplex (Hybrid I)

b - 5 refers to 5-S-DNP-dU

Attachment of a Crosslinking Group at a Specific Position in DNA

The analog-modified oligonucleotide (100 pmol) was radiolabeled with T4 polynucleotide kinase using $[\gamma^{32}P]$ ATP (Figure 5, Step 1), isolated by ethanol precipitation and re-dissolved in water. The presence of the 5-S-DNP group on the 5' terminal nucleotide did not interfere with recognition and modification of the 5' OH group by kinase. The oligonucleotide was then incorporated into double-stranded DNA by either PCR (not shown) or by primer-extension. For primer-extension using a single-stranded phagemid DNA template, the oligonucleotide (20 pmoles) was annealed to the template (1 pmol) by heating the mixture to 75°C, cooling to 65°C and then placing the mixture on ice (Step 2). The oligonucleotide was then extended with a

thermostable DNA polymerase (AmpliTaq, Perkin-Elmer) by
incubating for 5 minutes at room temperature and then two hours
at 70°C. The template was ligated with T4 DNA ligase for 1 hour
at 37°C, and then ligated DNA was isolated by ethanol

5   precipitation (Step 3). Before addition of a crosslinking
group, the DNA protecting group must be removed (Step 4). This
was accomplished by treatment of the oligonucleotide (4.5 μL)
with 1.4 M β-mercaptoethanol either at 45°C for 4 hours or at
room temperature for 12 hours in 20 μL 20 mM Tris-HCl, pH 8.2.

10   The deprotected DNA was purified to remove the β-mercaptoethanol
by ethanol precipitation, and the DNA was dissolved in 20 μL
water. A photocrosslinking group was added to the DNA by
reaction with the alkylating agent azidophenacyl bromide, and
the DNA was again purified for transcription and crosslinking

15   (Step 5).

Photocrosslinking

     For UV crosslinking, the reactions are split in half; a
control aliquot is kept in the dark at room temperature while
the other aliquot is irradiated at room temperature for 2

20   minutes in a polystyrene tube 1.5 cm from a 302 nm light source
(Spectroline model XX-15B, 1800 μW/cm^2 at 15 cm). After
irradiation, DTT (to 60 mM) is added to all samples, and they
are left in the dark for at least 5 minutes. After irradiation,
but before protein gel electrophoresis, aliquots from the

25   irradiated chase reactions are treated with nuclease to digest
the DNA to a small radioactive fragment (Step 6). The nuclease
treatment results in digestion of the DNA to small radioactive
fragments. This leaves a small radioactive piece of DNA on
proteins which have been crosslinked to the DNA during

30   irradiation. The control and irradiated samples are then
analyzed by gel electrophoresis and autoradiography.

Electrophoretic Analysis of Crosslinked Proteins

     For identification of crosslinked proteins, samples are
mixed with an equal volume of 60 mM TrisCl (pH 8.0), 60 mM DTT,

35   3.4% (w/v) SDS, 17% (v/v) glycerol, 0.02% (w/v) bromophenol
blue, 0.02% (w/v) xylene cyanol, heated for 3 minutes at 94°C,

SUBSTITUTE SHEET (RULE 26)

and then analyzed on 12 cm x 0.75 cm 10% SDS/polyacrylamide gels
with a 4% stacking gel (acrylamide/methylene bisacrylamide =
27/1). Proteins are electrophoretically blotted onto MSI
NitroBind Nitrocellulose membrane (Westboro, MA) in 25 mM Tris,
192 mM glycine, 20% (v/v) methanol (Figure 6A). The membrane is
then silver-stained in 2% (w/v) sodium citrate, 0.8% (w/v)
ferrous sulfate, 0.1% (w/v) silver nitrate for 5 minutes and
dried. Autoradiography of the membranes is at -80°C with
Amersham Hyperfilm-MP and a Cronex Lightning Plus intensifying
screen.

5-S-DNP-dU can be used for site-specific modification of
DNA at internal and terminal positions with DNA, and molecular
probes can be placed at variable distances from the DNA
backbone. In addition, the commercial availability of
antibodies to the DNP group allows the direct use of the DNP-
labeled oligonucleotides as immunodetection probes.[15]

Example 4:   Utility: Assays for Diagnosis of Bladder Cancer

Bladder Cancer strikes over 50,000 Americans every year,
killing more than 11,000 (23). Currently used means to diagnose
bladder cancer usually rely on immunocytochemistry, nuclear DNA
content and cytology. Extensive use of highly trained labor
that these tests consume make them fairly expensive to
administer. In addition, while being fairly specific for the
diagnosis of bladder cancer, these tests lack sensitivity.
Furthermore, they are poor prognosticators of cancer
progression. Tests or biopsies with better sensitivity and
prognostic value are invasive, even more expensive and poorly
tolerated by patients. These facts result in many patients
being under-diagnosed or under-treated for their illnesses.

An ideal test would be sensitive, specific and prognostic
for bladder cancer, less invasive, require fewer highly skilled
scientists to administer, better tolerated by patients and less
costly. Such an assay would be expected to decrease mortality
and morbidity due to this disease. Such a test would also be
expected to represent a significant commercial possibility since
it would not only replace many of the existing tests in the
marketplace but would enable an increase in the number of tests

performed.

One of the methods currently used for detection of mRNA levels in a cell utilizes an RNase protection assay (Hershey and krause, 1989) which is commercially available. In an RNase

5   protection assay, a labeled RNA probe of discrete length is constructed that includes two domains (Figure 6). One domain compliments, and therefore hybridizes to, an mRNA species whose abundance one wishes to detect and quantify. This domain forms a double stranded RNA molecule with the intended mRNA target.

10   The other domain is made up of sequence that does not hybridize to the intended target. In an RNase protection assay, total cell RNA (or total cellular mRNA) is incubated with an excess of an RNA probe under conditions that favor specific hybridization of the probe with the mRNA species to be detected and

15   quantified. After the hybridization step, the RNA is digested with a mixture of RNases. All single stranded RNA is hydrolyzed down to nucleotides or short oligonucleotides. Only double stranded RNA is protected from this digestion. All of the unhybridized probe is degraded. This includes the domains of

20   the hybridizing probes that do not complement their targets.

The digested RNAs are separated according to size by gel electrophoresis and then visualized by autoradiography. A positive signal is a labeled RNA molecule that is the correct length to be that portion of the probe that had hybridized to

25   the intended target. This shortened form of the original probe can not be created by other means. This is why false positives are not a technical concern. Also, with an excess of probe and under the appropriate hybridization conditions, nearly all of the intended target is hybridized to the nucleic acid probe

30   (NAP). This makes the intensity of the signal proportional to the original concentration of the target mRNA species.

The advantages of an RNase protection assay are that it is quantitative, reproducible and completely free of false positives derived from technical difficulties. This last

35   feature makes such an assay format appealing as a technological platform for diagnostic determinations of patient derived materials. The disadvantages of an RNase protection assay are that the probes are very tedious to make. They possess very

short half lives because of the radioactive isotopes (^{32}P) used
to produce the signal.   In addition, RNA is inherently less
stable than DNA, and some RNases can be extremely difficult to
inactivate.

5   The disadvantages of RNase protection are overcome by the
envisioned NAP Nuclease Protection Assays (NAP-protection
assays) employing the derivatizable nucleotide analogs herein.
The NAPs can be synthesized in large quantities, carefully
evaluated (quality control assays), and then stored for future
10   use.   There is an economy of scale that this affords which
overcomes the difficulty of probe construction and synthesis.
These NAPs will utilize non-isotopic means of detection,
permitting the NAPs to have extended shelf lives. Additionally,
the NAPs for different target mRNAs can be modified with
15   different reporter groups. This will permit the examination of
the expression of several genes simultaneously. This assay can
be modified to accept automation to reduce run times.

        Use of oligonucleotides or NAP probes which contain
modified nucleotide analogs has been previously reported (21-23,
20   26-29).   The chemical approach to preparation of site-
specifically modified DNA involves incorporation of a modified
nucleotide during the automated synthesis of an oligonucleotide.
The modified oligonucleotide can then be incorporated into
double-stranded DNA by PCR amplification or primer extension of
25   a single-stranded DNA.   The use of various modified
phosphoramidites as masked synthons for preparing derivatized
oligonucleotides at both terminal and internal sites has been
reported (Xu, Zheng and Swann, 1992; Maurizi and Ginsburg,
1986). These masked synthons allow incorporation of protected
30   functional groups such as amines (Nagamachi, et al. 1974),
carboxylic acids, thiols (Goodwin and Glick, 1993) and
thiocarbonyls. Once incorporated the analogs are deprotected
and modified post-synthetically.   Convertible nucleoside
phosphoramidites, mononomers containing leaving groups, have
35   also been used to incorporate crosslinking and photocrosslinking
groups by post-synthetic substitution of the leaving group.
Several companies sell reagents for the addition of molecular
tags to either of the two ends of a nucleic acid (5' or 3').

50

Some provide reagents which allow the incorporation of more than one analog at internal positions in the NAP, to allow an increase in the amount of signal produced by the NAP, there is a resultant perturbation of nucleic acid structure or decrease

5   in hybridization function (interaction with mRNA) of the NAP. Since the assay is based on specific NAP hybridization, with the current methods available, increasing the signal in a NAP must be balanced against the loss of hybridization capacity. However, the goal is always to obtain the strongest signal

10   possible, because this not only decreases data acquisition time, but increases the level of sensitivity of a given assay. This is clearly an important methodology, for which a commercial market already exists.

Preferred versions of the nucleotide analogs contemplated

15   herein have the following characteristics which make them preferable to other currently available nucleotide analogs for this assay:

    1.   The nucleotide modification is on a position of the base which is not required for normal RNA and DNA

20   basepairing, therefore specific hybridization is still possible.

    2.   The reactive group added to the nucleotide is not normally found in DNA or mRNA (S group), therefore reaction of the probe modified oligonucleotides or

25   NAPs with thiol-reactive alkylating agents gives virtually thiol-specific modification of the NAP. Analogs with reactive amino groups, which are modified by alkylating agents targeting amines, are less specifically modified. This is due to the presence of

30   exocyclic amine groups in the normal nucleotides of both DNA and RNA, and these can be modified by the alkylating reagents as well.

    3.   The reactivity of the aryl thiol group on 5-SH-UTP to alkylating agents is considerably higher than that of

35   alkyl thio ether derivatives ($pK_A=5.5$ vs 8,).

    4.   The linker attached to the nucleotide base contains all single-bonds (vs alkyl or alkynyl), thereby allowing free rotation of the probe in the region

close to the nucleic acid backbone. This is advantageous because the flexibility of the linker are allows the nucleotide analogs to adopt to different enzyme active sites (in DNA and RNA polymerases), making them generally good substrates for enzymatic incorporation into nucleic acids.

5. As assessed by several functional assays, analogs modified through the aryl thio group of 5-S-U derivatives do not disturb normal RNA secondary or tertiary structures.

Unlike other types of cancer, the primary tumor in bladder cancer sheds or exfoliates relatively large numbers of cells into the lumen of the bladder. These cells can be easily collected in urine or from bladder washes. In a void urine from a patient with a bladder tumor, there will be very few to about a million cancer cells. Bladder washes yield more cells than void urines. Currently, the company UroCor, Inc. and other firms provide services to diagnose disease. Bladder cancer cells collected in urine are examined for morphology, DNA content and for protein biomarkers as detected immunochemically. While these tests are not as informative as desired, it is clear that cancer cells can be collected from patient urines and that these cells are relatively intact with respect to morphology, nucleic acid and protein content, and expression. RNA has not been directly examined in these cells, but it is reasonable to expect that the RNA content of these cells will also be intact relative to the needs of the envisioned NAP-protection assay.

A NAP-protection assay can be designed to determine the relative abundance of mRNAs for four surrogate biomarkers of bladder cancer and/or its progression relative to the abundance of a mRNA encoding a "house keeping" gene that is not differentially regulated. Changes in the abundance of the surrogate biomarkers relative to the undifferentially expressed gene will be informative in both diagnosing and evaluating the state of a patient's bladder cancer. In one embodiment, the final assay uses five probes, each specific for a different gene. One of the probes will hybridize to an appropriate "house keeping" gene. The hybridization signal from this probe will

52

act as one of the positive controls in the assay. Of the
remaining four mRNA species, targets will be selected such that
in cancerous cells two of the mRNAs are increased in abundance
and two are decreased relative to nontransformed cells. In this
5   format, each regulatory pathway acts as an internal control for
the other.

The following assays can be carried out to probe for mRNA
levels of the human autocrine motility factor receptor, which is
differentially expressed in bladder cancer cells. The DNA
10 sequence (SEQ. ID. NO:4) for this gene is shown in Figure 7.

Ribonuclease NAP Protection Assay

One ribonuclease assay is based upon the currently utilized
radioisotopic RNase protection assay, described above. However,
the radioactively labeled nucleotide substrates are replaced by
15 5-S-R-XTP nucleotides, where X is U, dU, C, or dC and R is one
of several different reporter groups. The NAPs for the assay
are ssRNA, which are synthesized enzymatically utilizing a
bacterial or phage RNA polymerase.

In this assay, RNA NAPs can be prepared in which the RNA is
20 labeled with many fluorescent, colormetric, chemiluminescent, or
antigenic reporter groups. One such antigenic reporter group is
the dinitrophenyl group (DNP), and the analog for these
experiments is proposed below. The assay is the same as the
current method previously described, but the protected RNA
25 products are detected by either monitoring fluorescent emission,
probing with antibodies, or monitoring chemiluminescence. In
this assay, final detection still requires the analysis of the
NAP protection assay by gel electrophoretic separation of the
products. However, the probes can be synthesized in large
30 quantities and stored, so this assay still has a major advantage
over the existing methodologies.

Deoxyribonuclease NAP Protection Assay

An alternative nuclease protection assay can utilize DNA
NAPs, similarly modified with multiple reporter groups. One
35 compelling reason to consider DNA NAPs is that they are both
chemically and biologically more stable than the comparable RNA

SUBSTITUTE SHEET (RULE 26)

NAPs. This makes them better prospects for synthesis in large quantities and for long storage. Despite this inherent stability relative to RNA, ssDNA can be easily degraded by single-strand specific DNases, in the manner as utilized for ssRNA digestion. In addition, with DNA NAPs, one can synthesize large quantities of the DNA probe by utilizing the DNA amplification method, PCR (polymerase chain reaction).

1)   DNA NAPs Prepared with Analog-Tagged Oligonucleotide
         Primers

Already developed for use in such a DNA NAP-protection assay is the analog 5-DNP-dU phosphoramidite. This analog can be incorporated chemically into ssDNA oligonucleotides utilizing automated synthesis. Therefore oligonucleotide primers for the PCR reaction can be made which contain multiple analogs, with nearly every "T" in the oligonucleotide synthesis and isolation, and we have shown that it is also retained during the thermal cycling reaction of PCR. Therefore, one can prepare DNA NAPs which contain several reactive thiol groups in the 5' region of the NAP, and these can then be deprotected and modified with a variety of alkylating agents. One can therefore prepare probes for several different mRNA targets, each with a different reporter group, and differentially quantify the levels of each in one reaction. This aspect of the method becomes important for the progression of the assay to automation.

2)   DNA NAPs Prepared with Analog-Tagged Deoxynucleoside
         Triphosphates

DNA NAPs can be synthesized by incorporating several (e.g., five to ten) 5-S-DNP-dUs into the 5'-tagged oligonucleotide, and then synthesizing full length DNA NAP by PCR with unmodified dNTPs, e.g., 5-S-DNP-dUTP.

Automation of NAP Protection Assays

Automation of either the RNA NAP Protection Assay or the DNA NAP Protection Assay requires that the electrophoresis step be eliminated. The separation of the individual protected mRNA species by electrophoresis is one of the more time-consuming

steps of the current assay, and makes this method intractable to automation. In order to eliminate this step, one must devise a method for the detection and relative quantitation of more than one mRNA species in a single reaction. Currently, the best

5 approach for this involves the use of a variety of fluorescent tags, each with different excitation and emission properties. An increasing number of alkylating reagents tagged with fluorescent groups such as fluorescein or rhodamine, to name only two of many, are now commercially available. Companies are

10 designing fluorescent detectors which can differentiate among the different signals produced by these fluorescent dyes. This approach is modeled upon the automated assays currently used for nonisotopic DNA sequencing. The final choice of fluorescent dye combinations used for these assays are determined empirically

15 during the course of the first year of this grant. Alternatively, one can label different NAPs with different antigenic groups, so that RNAs could be distinguished with antibodies, a method already utilized in many automated assays.

Deoxyribonuclease NAP Protection Assay

20   For the DNA NAP assay, one can incorporate commonly used methods for the immobilization of DNA onto resins or microtiter plates. If, for example, a 96-well microtiter plate were coated with streptavidin, then DNA tagged at the 5' end with biotin could be efficiently immobilized. In this way, the DNA NAP and

25 any RNA to which it hybridizes, could be immobilized on a solid matrix, while other components of the hybridization mixture could be washed away.
        The following approach can be followed for establishment of an automated DNase NAP Protection Assay.

30   1)   DNA oligonucleotides can be synthesized which are tagged at the 5' end with one of many groups that would allow immobilization of the DNA to a solid matrix (for example biotinylated nucleotides and streptavidin-agarose resin). The DNA NAPs can be multiply-labeled by either:

35   a)   incorporating several (5-10) 5-S-DNP-dUs into the
             5'-tagged oligonucleotide, and then synthesizing

SUBSTITUTE SHEET (RULE 26)

full length DNA NAP by PCR with unmodified dNTPs.

b)   incorporating multiple (50-100) 5-S-DNP-dUs into the
body of the DNA product produced by PCR by using 5-
S-DNP-dUTP as a substrate.

5   2)   The 5' end-labeled, internally reporter-labeled DNA
NAPs can be deprotected and modified with one of the reporter
groups chosen in the studies described above.

3)   Reaction mixtures can be prepared containing multiple
mRNA targets and one or more reporter-tagged NAPs.

10   4)   After hybridization, reactions mixtures can be treated
with nucleases specific for single-stranded nucleic acids (such
as S1 or Mung Bean Nucleases), to produce the protected mRNA
species.

5)   Reactions mixtures can be added to microtiter plates
15   coated with Streptavidin (for 5'-biotin labeled DNA) or other
materials which are specific for binding the groups commonly
used to tag DNA probes. Since the 5' end of the DNA is
protected during the NAP-protection assay (see Fig. 1), and NAP
which does not   hybridize are degraded by nuclease, the only
20   material which   should bind to the solid matrix is the DNA-
mRNA hybrid.

6)   Microtiter   plates   can   be   washed   to   remove
unincorporated   nucleotides, thereby removing essentially all
background signal.

25   7)   Plates can be analyzed with a fluorescent microtiter
plate reader to quantify the relative amounts of mRNAs produced
for different NAPs.

56

REFERENCES

Agrawal, S., ed. <u>Methods in Molecular Biology</u>, ed. Walker, J.M., Vol. 20, Humana Press: Totowa, New Jersey (1993).

Barone, A.D., Tang, J.-Y. and Caruthers, M.H. (1984) *In situ* Activiation of bis-Dialkylaminophosphines - A New Method for Synthesizing Deoxyoligonucleotides on Polymer Supports *Nucleic Acids Res. 12*, 4051-4061.

Bartholomew, B., Braun. F.R., Kassavetis, G.A. and Geiduschek, E.P. (1994) Probing Close DNA Contacts of RNA Polymerase III Transcription Complexes with the Photoactive Nucleoside 4-Thiodeoxythymidine *J. Biol. Chem. 269*, 18090-18095.

Bartholomew, B., Kassavetis. G.A., Braun, B.R. and Geiduschek, E.P. (1990) The Subunit Structure of *Saccharomyces cerevisiae* Transcription Factor IIIC Probed with a Novel Photocrosslinking Reagent *EMBO J. 9*, 2197-2205.

Beaucage, S.L. and Iyer, R.P. (1993) The Functionalization of Oligonucleotides Via Phosphoramidite Derivatives *Tetrahedron 49*, 1925-1963.

Bergstrom, D.E., Beal, P., Jenson, J. and Lin, X. (1991) Palladium-mediated Synthesis of C-5 Pyrimidine Nucleoside Thioethers from Disulfides and Mercurinucleosides *J. Org. Chem. 56*, 5598-5602.

Bleasdale, C., Ellwood, S.B. and Golding, B.T. (1990) 4,4'-Dimethoxytrityl and 4-Monomethoxytrityl Tetrafluoroborate: Convenient Reagents for the Protection of Primary Alcohols Including Sugars *J. Chem. Soc. Perkin Trans 1* 803-804.

Bradley, D.H. and Hanna, M.M. (1992) Synthesis and Utility of 5-
     Thiocyanato Dexoyuridine and Uridine Phosphoramidites as
     Masked Synthons *Tetrahedron Lett.* 33, 6223-6226.

Brown, H.C., Narasimhan, S. and Choi, Y.M. (1982) Selective
     Reduction. 30. Effect of Cation and Solvent on the
     Reactivity of Saline Borohydrides for Reduction of
     Carboxylic Esters. Improved Procedures for the Conversion
     of Esters to Alcohols by Metal Borohydrides *J. Org. Chem.*
     47, 4702-4708.

Chaudhary, S.K. and Hernandez, O. (1979) A Simplified Procedure
     for the Preparation of Triphenylmethyleters *Tetrahedron
     Lett.* 95-98.

Christopherson, M.S. and Broom, A.D. (1991) Synthesis of
     Oligonucleotides Containing 2'-Deoxy-6-thioguanosine at a
     Predetermined Site *Nucleic Acids Res.* 19, 5719-5724.

Clegg, R.M., Murchie, A.I.H., Zechel, A., Carlberg, C.,
     Diekmann, S. and Lilley, D.M.J. (1992) Fluorescence
     Resonance Energy Transfer Analysis of the Structure of the
     Four-Way DNA Junction *Biochemistry 31*, 4846-4856.

Clivio, P., Fourrey, J.-L., Gasche, J., Audic, A., Favre, A.,
     Perrin, C.; and Woisard, A. (1992) Synthesis and
     Purification of Oligonucleotides Containing Sulfur
     Substituted Nucleobase: 4-Thiouracil, 4-Thiothymine and 6-
     Mercaptopurine *Tetrahedron Lett.* 33, 65-68.

Coleman, R.S. and Kesicki, E.A. (1994) Synthesis and
     Postsynthetic Modification of Oligodeoxynucleotides
     Containing 4-Thio-2'-deoxyuridine (d^{s4}U) *J. Am. Chem. Soc.*
     116, 11636-11642.

Connolly, B.A. and Newman, P.C. (1989) Synthesis and Properties
     of Oligonucleotides Containing 4-Thiothymine, 5-Methy-2-
     pyrimidinone-1-β-D-(2'-deoxyriboside) and 2-Thiothymidine
     Nucleic Acids Res. 17, 4957-4974.

Cosstick, R. and Douglas, M.E. (1991) Synthesis of a
     Dinucleoside Monophosphate Analogue Containing 6-*N*-(2-
     Aminoethyl)-2'-Deoxyadenosine. A Novel Approach to
     Sequence Specific Crosslinking in Synthetic
     Oligonucleotides *J. Chem. Soc. Perkin Trans.* 1 1035-1040.

Dissinger, S. and Hanna, M.M., (1990) <u>J. Biol. Chem.</u>, 265:7662.

SUBSTITUTE SHEET (RULE 26)

Flaschka, H.A., Barnard, A.J. and Sturrock, P.E. (1969) *Quantitative Analytical Chemistry, Short Introduction to Practice*, Barnes & Noble, INC., New York.

Gait, M.J., ed. <u>Practical Approach Series</u>, ed. Rickwood, D. and Harnes, B.D., IRL Press: Oxford (1984).

Gao, H., Fathi, R., Gaffney, B.L., Goswami, B., Kung, P.-P., Rhee, Y., Jin, R. and Jones, R.A. (1992) 6-*O*-(Pentatluorophenyl)-2'-deoxyguanosine: A Versatile Synthon for Nucleoside and Oligonucleotide Synthesis *J. Org. Chem.* 57, 6954-6959.

Glick, G.D. (1991) Synthesis of a Conformationally Restricted DNA Hairpin *J. Org. Chem.* 56, 6746-6747.

Gibson, K.J. and Benkovic, S.J. (1987) <u>Nucleic Acid Res.</u>, 15, No. 16:6455.

Goodwin, J.T. and Glick, G.D. (1993) Incorporation of Alkylthiol Chains at C-5 of Deoxyuridine *Tetrahedron Lett.* 34, 5549-5552.

Goodwin, J.T., Osborne, S.E., Swanson, P.C. and Glick, G.D. (1994) Synthesis of a Disulfide Cross-linked DNA Triple Helix *Tetrahedron Lett.* 35, 4527-4530.

Greene, T.W. (1981) *Protective Groups in Organic Synthesis*, John Wiley & Sons, New York.

Grzybowski, J., Will, D.W., Randall, R.E., Smith, C.A. and Brown, T. (1993) Synthesis and Antibody-mediated Detection of Oligonucleotides Containing Multiple 2,4-Dinitrophenyl Reporter Groups *Nucleic Acids Res.* 21, 1705-1712.

Hanna, M.M. (1989) <u>Methods in Enzymology</u>, 180:383

Hanna, M.M. (1993) Photocrosslinking Analysis of Protein-RNA Interactions in *E. coli* Transcription Complexes *Cellular Molecular Biology Research* 39, 393-399.

Hanna, M.M., Dissinger, S., Williams, B.D. and Colston, J.E. (1989) Synthesis and Characterization of 5-[4-(Azidophenacyl)thio]uridine 5'-Triphosphate, a Cleavable Photo-Cross-Linking Nucleotide Analogue *Biochemistry 28*, 5814-5820.

He, B. and Hanna, M.M., (1995) <u>Nucleic Acids Res.</u>, 23.

Hershey, A.D. and Krause, J.E., (1989) <u>Science</u>, 247:960.

Jadhav, V., Barawkar, D., Natu, A., and Ganesh, K. *Nucleosides*

and Nucleotides, (1997) 16, 107-114.

Kalman, T.I. and Bardos, T.J. (1967) The Autoxidation of 5-Mercaptouracil and 5-Mercaptodeoxyuridine *J. Am. Chem. Soc.* *89*, 1171-1175.

Keller, G.H., and Manak, M.M., (1989) <u>DNA Probes</u>, New York, N.Y.: Stockton Press.

Kini, G.D., Robins, R.K. and Avery, T.L. (1989) Synthesis and Antitumor Activity of Ribavirin Imidates. A New Facile Synthesis of Ribavirin Amidine (1-β-D-Ribofuranosyl-1,2,4-triazone-3-carboxamidine Hydrochloride) *J. Med. Chem.* *32*, 1447-1449.

Kricka, L.J. (1992) ed. <u>Nonisotopic DNA Probe Techniques</u>, Academic Press, Inc.: San Diego, CA.

Kuimelis, R.G. and Nambiar, K.P. (1994) Synthesis of Oligodeoxynucleotides Containing 2-Thiopyrimidine Residues-A New Protection Scheme *Nucleic Acids Res.* *22*, 1429-1436.

Lakshman, M.K., Sayer, J.M., Yagi, H. and Jerina, D.M. (1992) Synthesis and Duplex-Forming Properties of a Nonanucleotide Containing an N^6-Deoxyadenosine Adduct of a Bay-Region Diol Epoxide *J. Org. Chem.* *57*, 4585-4590.

Maurizi, M.R. and Ginsburg, A. (1986) <u>Biochemistry</u>, 25:131.

Mergny, J.-L., Boutorine, A.A., Garestier, T., Belloc, F., Rougee, M., Bulychev, N.V., Koshkin, A.A., Bourson, J., Lebedev, A.V., Valeur, B., Thuong, N.T. and Helene, C. (1994) Fluorescence energy transfer as a probe for nucleic acid structures and sequences *Nucleic Acids Res.* *22*, 920,928.

Nagamachi, T., Fourrey, J.-L., Torrence, P.F., Waters, J.A. and Witkop, B. (1974) Synthesis, Chemistry, and Biological Activity of 5-Thiocyanatopyrimidine Nucleosides as Potential Masked Thiols *J. Med. Chem.* *17*, 403-406.

Scremin, C.L., Zhou, L., Srinivasachar, K. and Beaucage, S.L. (1994) Stepwise Regeneration and Recovery of Deoxyribonucleoside Phosphoramidite Monomers during Solid-Phase Oligonucleotide Synthesis *J. Org. Chem.* *59*, 1963-1966.

Shaltiel, S. (1967) Thiolysis of Some Dinitrophenyl Derivatives of Amino Acids *Biochem. Biophys. Res. Comm.* *29*, 178-183.

SUBSTITUTE SHEET (RULE 26)

Sheardy, R.D. and Seeman, N.C. (1986) The Synthesis of a
     Deoxyoligonucleotide Incorporating 5-Iododeoxyruidine *J.*
     Org. Chem. 51, 4301-4303.

Sinha, N.D., Biernat, J., McManus, J., and Koster, H. (1984)
     Polymer Support Oligonucleotide Synthesis XVIII: Use of β-
     cyanoethyl-*N,N*-dialkylamino-/*N*-morpholino Phosphoramidite
     of Deoxynucleosides for the Synthesis of DNA Fragments
     Simplifying Deprotection and Isolation of the Final Product
     Nucleic Acids Res. 12, 4539-4557.

Soai, K. and Ookawa, A. (1986) Mixed Solvents Containing
     Methanol as Useful Reaction Media for Unique Chemoselective
     Reductions with Lithium Borohydride *J. Org. Chem. 51*, 4000-
     4005.

Torrence, P.F., Waters, J.A. and Witkop, B. (1968) 2'Deoxy-
     5'(thiocyanato)uridine [1-(2-deoxy-β-D-
     erythropentofuranosyl)-5-(thiocyanato)uracil] *Synthetic*
     Procedures in Nucleic Acid Chemistry, (W.W. Zorbach and
     R.S. Tipson, Ed.) pp 367-370, Interscience, New York.

Trivedi, B.K. and Bruns, R.F. (1989) C2.N^6-Disubstituted
     Adenosines: Synthesis and Structure-Activity Relationships
     J. Med. Chem. 32, 1667-1673.

Waters, T.R. and Connolly B.A. (1992) Straightforward Synthesis
     of 6-Thiodeoxyguanosine and Its Incorporation into
     Oligodeoxynucleotides *Nucleosides Nucleotides 11*, 985-998.

Xu, Y.-Z., Zheng, Q. and Swann, P.F. (1992a) Synthesis and
     Duplex Stability of Oligodeoxynucleotides Containing 6-
     Mercaptopurine *Tetrahedron Lett. 33*, 5837-5840.

Xu, Y.-Z., Zheng, Q. and Swann, P.F. (1992b) Synthesis by Post-
     Synthetic Substitution of Oligomers Containing Guanine
     Modified at the 6-Position with S-, N-, O-derivatives
     Tetrahedron 48, 1729-1740.

Xu, Y.-Z., Zheng, Q. and Swann, P.F. (1992c) Synthesis of DNA
     Containing Modified Bases by Postsynthetic Substitution.
     Synthesis of Oligomers Containing 4-Substituted Thymine:
     O^4-Alkylthymine, 5-Methylcytosine, *N*^{4}-(Dimethylamino)-5-
     methylcytosine, and 4-Thiothymine *J. Org. Chem. 57*, 3839-
     3845.

Yang, X.J., Hart, C.M., Grayhack, E.J. and Roberts, J.W. (1987)

SUBSTITUTE SHEET (RULE 26)

61

Transcription Antitermination by Phage Lambda Gene Q Protein Requires a DNA Segment Spanning the RNA Start Site *Genes Dev 1*, 217-226.

5   Zhang, Y. and Hanna, M.M. (1995) Expression and Functional Characterization of *Escherichia coli* NusA and Lambda Q as Glutathione-S-Transferase Fusion Proteins *Protein Expression and Purification 6*, 625-631.

62

Changes may be made in the compounds described herein or in the steps or the sequence of steps of the methods of synthesis or use of the compounds described herein without departing from the spirit and scope of the invention as defined in the following claims.

What is claimed is:

    1.   A nucleotide analog of the formula:

wherein

    R_1 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

    R_2 is -H, -OH, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, -OR$_4$, or a solid support bonded via an O at the 3' position;

    R_3 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

    R_4 is a lower alkyl or a protecting group; and

    B is a modified pyrimidine base comprising a protected thiol group attached at the 5 position on said base that is not involved in Watson-Crick base pairing or does not disrupt normal Watson-Crick base pairing, said protected thiol group being stable under conditions of chemical nucleic acid synthesis and/or conditions of enyzmatic nucleic synthesis and being convertible to a reactive thiol after said synthesis.

    2.   The nucleotide analog of Claim 1 wherein said base is selected from the group consisting of cytosine and uracil.

    3.   The nucleotide analog of Claim 1 wherein said base compromises an additional protecting group on a reactive moiety of said base.

    4.   The nucleotide analog of Claim 3 wherein said base is selected from the group consisting of N^4-anisoyl cytosine, N^4-

64

benzoyl cytosine, N⁴-isobutyryl cytosine and N⁴-acetyl cytosine.

5.   The nucleotide analog of Claim 1 wherein said protected thiol group on said base is selected from the group consisting of thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R_{10}-DNP), and alkyldisulfide (-S-S-R_{10}) wherein R_{10} is a lower alkyl.

6.   The nucleotide analog of Claim 1, wherein the protecting group of R_1 is selected from the group consisting of p-(dimethoxytrityl), p-(monomethoxytrityl), fluorenylmethyl-oxycarbonyl, levuloyl and 9-phenylxanthene-9-yl.

7.   The nucleotide analog of Claim 1, wherein protecting group of R_3 is selected from the group consisting of 1-(2-chloro-4-methylphenyl)-4-methoxy-4-piperidinium, 2'-acetal, o-nitrobenzyl, tert-butyldimethyl silyl, tetrahydrofuranyl and 4-methoxytetrahydropyranyl.

8.   The nucleotide analog of Claim 1 wherein R_1 is -OR_4 wherein R_4 is a protecting group and R_2 is a phosphoramidite group.

9.   The nucleotide analog of Claim 8 wherein said base is cytosine, or uracil.

10.   The nucleotide analog of Claim 9 wherein said base comprises an additional protecting group on a reactive moiety of said base.

11.   The nucleotide analog of Claim 12 wherein said base is selected from the group consisting of N⁴-anisoyl cytosine, N⁴-benzoyl cytosine, N⁴-isobutyryl cytosine and N⁴-acetyl cytosine.

12.   The nucleotide analog of Claim 8 wherein said protected thiol group of said base is selected from the group consisting of thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R_{10}-DNP), and alkyldisulfide (-S-S-R_{10}) wherein R_{10} is a lower alkyl.

SUBSTITUTE SHEET (RULE 26)

13.   The nucleotide analog of Claim 8 wherein the protecting group of R_1 is selected from the group consisting of p-(dimethoxytrityl), p-(monomethoxytrityl), fluorenylmethyl-oxycarbonyl, levuloyl, and 9-phenylxanthene-9-yl.
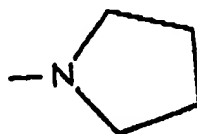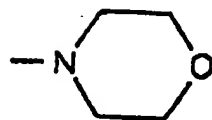
14.   The nucleotide analog of Claim 8, wherein the protecting group of R_3 is selected from the group consisting of 1-(2-chloro-4-methylphenyl)-4-methoxy-4-piperidinyl, 2'-acetal, o-nitrobenzyl, tert-butyldimethyl silyl, tetrahydrofuranyl and 4-methoxytetrahydropyranyl.

15.   The nucleotide analog of Claim 1 or 8, wherein said phosphoramidite of R_2 is represented by the formula:

R_6 is a lower alkyl, cyanoethyl or a substituted lower alkyl; and

R_7 and R_8 are independently lower alkyls, or when taken together with the nitrogen to which they are attached comprise one of the groups:
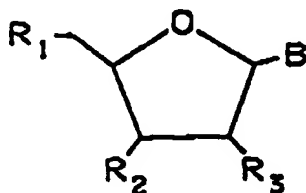

16.   The nucleotide analog of Claim 1 wherein R_1 is $-OR_4$ wherein R_4 is a protecting group and R_2 is a phosphorothioamidite group.

17.   The nucleotide analog of Claim 1 or 16, wherein said R$_2$ is a phosphorothioamidite represented by the formula:

wherein

R$_6$ is a lower alkyl, cyanoethyl or a substituted lower
       alkyl; and

R$_7$ and R$_8$ are independently lower alkyls, or when taken
       together with the nitrogen to which they are
       attached comprise one of the groups:

18.   The nucleotide analog of Claim 1 wherein R$_2$ is an O-substituted monophosphate group selected from the group consisting of -2-chlorophenyl monophosphate, O-2,5-dichlorophenyl monophosphate, O-2,2,2-trichloroethyl monophosphate and the N oxide of 4-methoxypyridine-2-methylene monophosphate.

19.   A nucleotide analog of the formula:

wherein

R$_1$ is -H, -OH, a mono, di, or triphosphate group, or
      -OR$_4$;

R$_2$ is a phosphoramidite group;

SUBSTITUTE SHEET (RULE 26)

R_3 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R_4 is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a thiodinitrophenyl group attached at the 5 position on said pyrimidine base.

20.   The nucleotide analog of Claim 19 wherein said base is selected from the group consisting of N^4-anisoyl cytosine, N^4-benzoyl cytosine, N^4-isobutyryl cytosine and N^4-acetyl cytosine.

21.   The nucleotide analog of Claim 19, wherein R_1 is -OR$_4$ and R_4 is a protecting group selected from the group consisting of p-(dimethoxytrityl), p-(monomethoxytrityl), fluorenylmethyl-oxycarbonyl, levuloyl, and 9-phenylxanthene-9-yl.

22.   The nucleotide analog of Claim 19 wherein said phosphoramidite of R_2 is represented by the formula:

wherein

R_6 is a lower alkyl, cyanoethyl or substituted lower alkyl; and

R_7 and R_8 are independently lower alkyls, or when taken together with the nitrogen to which they are attached form the groups:

23.   The nucleotide analog of Claim 22 wherein R_6 is 2-cyanoethyl, R_7 is isopropyl and R_8 is isopropyl.

SUBSTITUTE SHEET (RULE 26)

68

24.   The nucleotide analog of Claim 23 wherein said base is uridine or deoxyuridine.

25.   The nucleotide analog of Claim 24 wherein R_1 is $-OR_4$ and R_4 is p-(dimethoxytrityl).

26.   A nucleotide analog of the formula:

wherein:

R_1 is $-H$, $-OH$, a mono, di, or triphosphate group, or $-OR_4$;

5   R_2 is $-H$, $-OH$, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, $-OR_4$, or a solid support bonded via an O at the 3' position;

10   R_3 is $-H$, $-OH$, a mono, di, or triphosphate group, or $-OR_4$;

R_4 is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at the 5 position on said

15   base.

27.   The nucleotide analog of Claim 26 wherein said base is cytosine, or uracil.

28.   The nucleotide analog of Claim 26 wherein said protected thiol group on said base is selected from the group consisting of thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R_{10}-DNP), S-phenylacetamidomethyl(-S-CH_2NHCOCH$_2$Ph) wherein R_{10} is a lower alkyl.

29.   The nucleotide analog of Claim 26, wherein R_1 is a monophosphate, a diphosphate, or a triphosphate.

SUBSTITUTE SHEET (RULE 26)

69

30. An oligonucleotide containing the nucleotide analog of any one of Claims 1, 8, 16, 19, or 26.

31. A method of producing the nucleotide analog of any one of Claims 1, 8, 16, 19, or 26 which comprises preparing a thiol-protected nucleoside or nucleotide base wherein said thiol is attached to the 5 position on said base; reacting said nucleoside or nucleotide base under conditions to effect conversion of said base to a phosphoramidite, phosphorothioamidite, phosphonate, O-substituted monophosphate or phosphate nucleotide analog and under conditions which do not destroy the protected thiol; and recovering said analog.

32. A method of synthesizing a nucleic acid having an attached functional group which comprises incorporating a thiol-protected nucleotide analog of any one of Claims 1, 8, 16, 19, or 26 into a nucleic acid by a chemical or enzymatic method of nucleic acid synthesis; recovering said nucleic acid containing said analog; deprotecting the analog of said nucleic acid to produce a nucleic acid containing a reactive thiol group; treating the reactive thiol group with a thiol modifying reagent to thereby attach a functional group and produce said nucleic acid with an attached functional group; and recovering said nucleic acid with said attached functional group.

33. The method of Claim 32 wherein said functional group is selected from the group consisting of a photocrosslinker, a crosslinker, a reporter molecule, a radioisotope, a fluorescent group, a spin label, chemiluminescent or an antigenic group.

34. The method of Claim 33 wherein said photocrosslinker is an aryl azide.

35. The method of Claim 33 wherein said reporter group is biotin, an enzyme, or a fluorescent molecule.

36. The method of Claim 33 wherein said fluorescent group is fluorescein.

SUBSTITUTE SHEET (RULE 26)

37. A method for producing a deoxyuridine analog having a
protected thiol group on the 5 position, comprising:

      reacting a diacetyl 5-thiocyanate deoxyuridine to form
        a 5-dinitrophenylthio deoxyuridine;

5   converting said 5-dinitrophenylthio deoxyuridine to a
        phosphoramidite, phosphorothioamidite,
        phosphonate, O-substituted monophosphate or
        phosphate nucleotide analog under conditions
        which do not destroy the 5-dinitrophenylthio

10   group;

     and recovering said 5-dinitrophenylthio deoxyuridine
        analog.

SUBSTITUTE SHEET (RULE 26)

methylphenyl)-4-methoxy-4-piperidinyl, 2'-acetal, o-nitrobenzyl, tert-butyldimethyl silyl, tetrahydrofuranyl and 4-methoxytetrahydropyranyl.
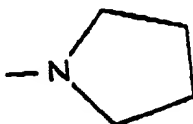
15. The nucleotide analog of Claim 1 or 8, wherein said phosphoramidite of R_2 is represented by the formula:

R_6 is a lower alkyl, cyanoethyl or a substituted lower alkyl; and

R_7 and R_8 are independently lower alkyls, or when taken together with the nitrogen to which they are attached comprise one of the groups:

 , , or

16. The nucleotide analog of Claim 1 wherein R_1 is $-OR_4$ wherein R_4 is a protecting group and R_2 is a phosphorothioamidite group.

-89-

17. The nucleotide analog of Claim 1 or 16, wherein said R_2 is a phosphorothioamidite represented by the formula:

$$\underset{R_6}{\overset{\overset{\displaystyle O}{|}}{\underset{S}{\underset{|}{P}}}}\!\!-\!\!N\!\!\begin{array}{c} R_7 \\ R_8 \end{array}$$
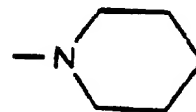
wherein

R_6 is a lower alkyl, cyanoethyl or a substituted lower alkyl; and

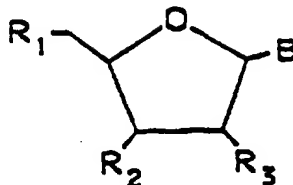R_7 and R_8 are independently lower alkyls, or when taken together with the nitrogen to which they are attached comprise one of the groups:

, , or

18. The nucleotide analog of Claim 1 wherein R_2 is an O-substituted monophosphate group selected from the group consisting of -2-chlorophenyl monophosphate, O-2,5-dichlorophenyl monophosphate, O-2,2,2-trichloroethyl monophosphate and the N oxide of 4-methoxypyridine-2-methylene monophosphate.

19.   A nucleotide analog of the formula:

wherein

R₁ is -H, -OH, a mono, di, or triphosphate group, or
       -OR₄;

5   R₂ is a phosphoramidite group;

R₃ is -H, -OH, a mono, di, or triphosphate group, or
       -OR₄;

R₄ is a lower alkyl or a protecting group; and

B   is   a   modified   pyrimidine   base   comprising   a
10   thiodinitrophenyl group attached at the 5 position
           on said pyrimidine base.

20.   The nucleotide analog of Claim 19 wherein said base is
selected from the group consisting of N⁴-anisoyl cytosine, N⁴-
benzoyl cytosine, N⁴-isobutyryl cytosine and N⁴-acetyl cytosine.

21.   The nucleotide analog of Claim 19, wherein R₁ is -OR₄ and
R₄ is a protecting group selected from the group consisting of p-
(dimethoxytrityl),   p-(monomethoxytrityl),   fluorenylmethyl-
oxycarbonyl, levuloyl, and 9-phenylxanthene-9-yl.

22.   The nucleotide analog of Claim 19 wherein said phosphoramidite of R_2 is represented by the formula:

$$O—P(O—R_6)—N(R_7)(R_8)$$

wherein

   R_6 is a lower alkyl, cyanoethyl or substituted lower alkyl; and

   R_7 and R_8 are independently lower alkyls, or when taken together with the nitrogen to which they are attached form the groups:

23.   The nucleotide analog of Claim 22 wherein R_6 is 2-cyanoethyl, R_7 is isopropyl and R_8 is isopropyl.

24.   The nucleotide analog of Claim 23 wherein said base is uridine or deoxyuridine.

25.   The nucleotide analog of Claim 24 wherein R_1 is $-OR_4$ and R_4 is p-(dimethoxytrityl).

26.   A nucleotide analog of the formula:

wherein:

R₁ is -H, -OH, a mono, di, or triphosphate group, or
-OR₄;

R₂ is -H, -OH, a mono, di, or triphosphate group, a
phosphoramidite group, a phosphorothioamidite
group, a phosphonate group, an O-substituted
monophosphate group, -OR₄, or a solid support bonded
via an O at the 3' position;

R₃ is -H, -OH, a mono, di, or triphosphate group, or
-OR₄;

R₄ is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected
thiol group attached at the 5 position on said
base.

27.   The nucleotide analog of Claim 26 wherein said base is
cytosine, or uracil.

28.   The nucleotide analog of Claim 26 wherein said protected
thiol group on said base is selected from the group consisting of
thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R₁₀-DNP), S-
phenylacetamidomethyl(-S-CH₂NHCOCH₂Ph) wherein R₁₀ is a lower alkyl.

-93-

29. The nucleotide analog of Claim 26, wherein R_1 is a monophosphate, a diphosphate, or a triphosphate.

30. An oligonucleotide containing the nucleotide analog of any one of Claims 1, 8, 16, 19, or 26.

31. A method of producing the nucleotide analog of any one of Claims 1, 8, 16, 19, or 26 which comprises preparing a thiol-protected nucleoside or nucleotide base wherein said thiol is attached to the 5 position on said base; reacting said nucleoside or nucleotide base under conditions to effect conversion of said base to a phosphoramidite, phosphorothioamidite, phosphonate, O-substituted monophosphate or phosphate nucleotide analog and under conditions which do not destroy the protected thiol; and recovering said analog.
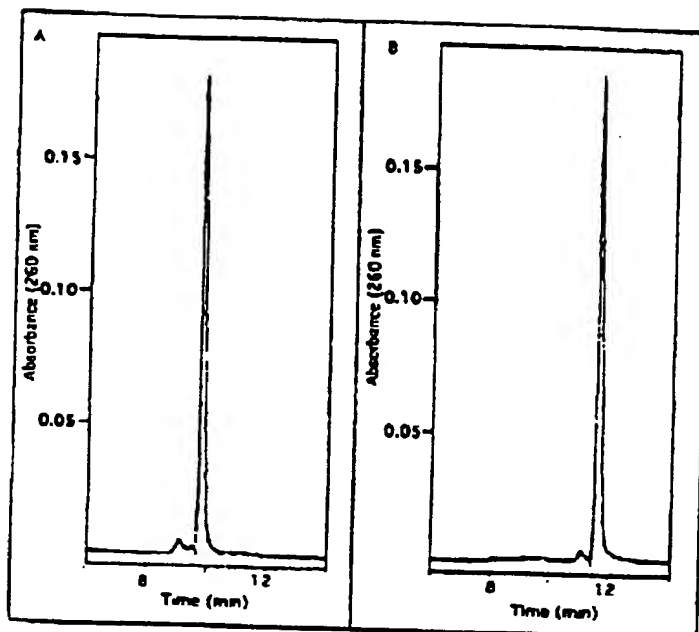

32. A method of synthesizing a nucleic acid having an attached functional group which comprises incorporating a thiol-protected nucleotide analog of any one of Claims 1, 8, 16, 19, or 26 into a nucleic acid by a chemical or enzymatic method of nucleic acid synthesis; recovering said nucleic acid containing said analog; deprotecting the analog of said nucleic acid to produce a nucleic acid containing a reactive thiol group; treating the reactive thiol group with a thiol modifying reagent to thereby attach a functional group and produce said nucleic acid with an

-94-

10   attached functional group; and recovering said nucleic acid with
said attached functional group.

33.   The method of Claim 32 wherein said functional group is
selected from the group consisting of a photocrosslinker, a
crosslinker, a reporter molecule, a radioisotope, a fluorescent
group, a spin label, chemiluminescent or an antigenic group.

34.   The method of Claim 33 wherein said photocrosslinker is
an aryl azide.

35.   The method of Claim 33 wherein said reporter group is
biotin, an enzyme, or a fluorescent molecule.

36.   The method of Claim 33 wherein said fluorescent group is
fluorescein.

37. A method for producing a deoxyuridine analog having a
protected thiol group on the 5 position, comprising:

        reacting a diacetyl 5-thiocyanate deoxyuridine to form a
            5-dinitrophenylthio deoxyuridine;

5   converting said 5-dinitrophenylthio deoxyuridine to a
            phosphoramidite, phosphorothioamidite, phosphonate,
            O-substituted monophosphate or phosphate nucleotide
            analog under conditions which do not destroy the 5-
            dinitrophenylthio group;

-95-

10

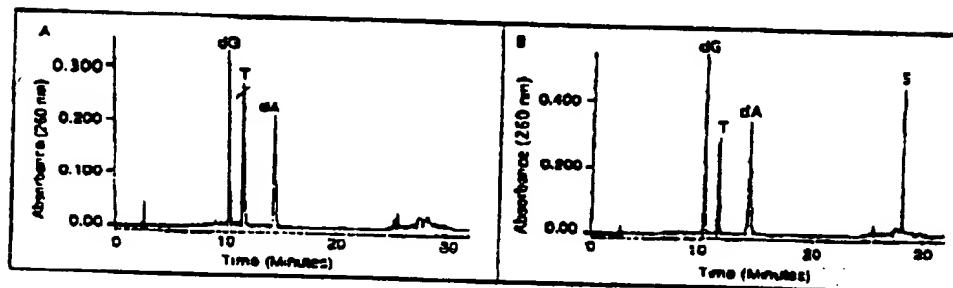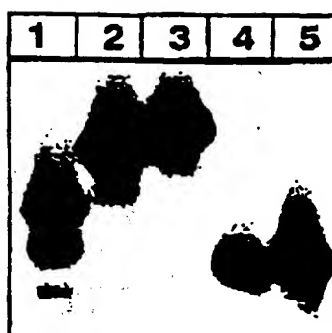and recovering said 5-dinitrophenylthio deoxyuridine analog.

A.

B.
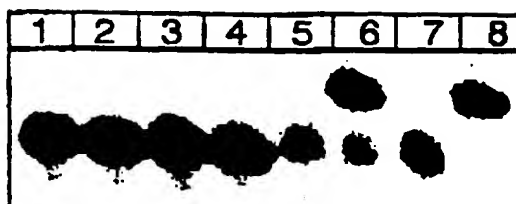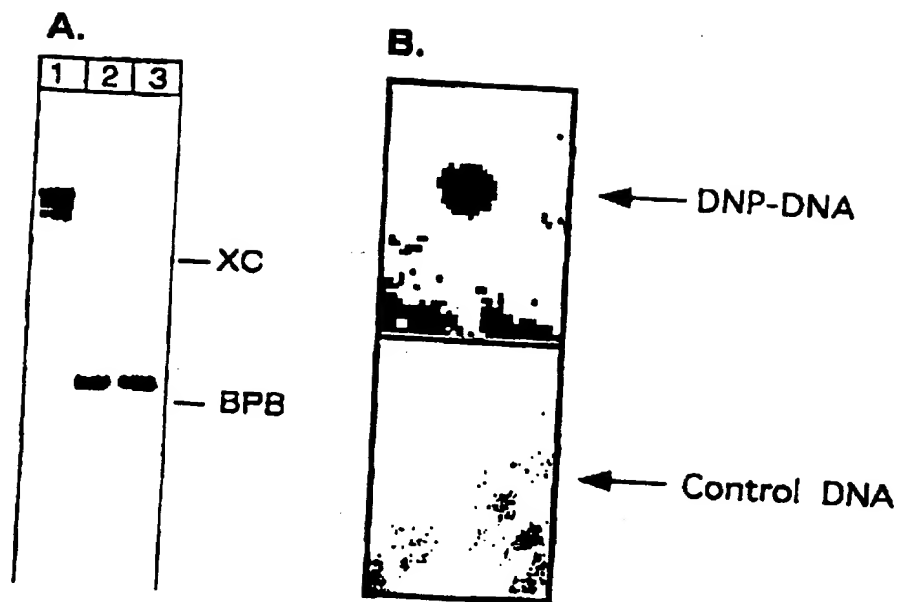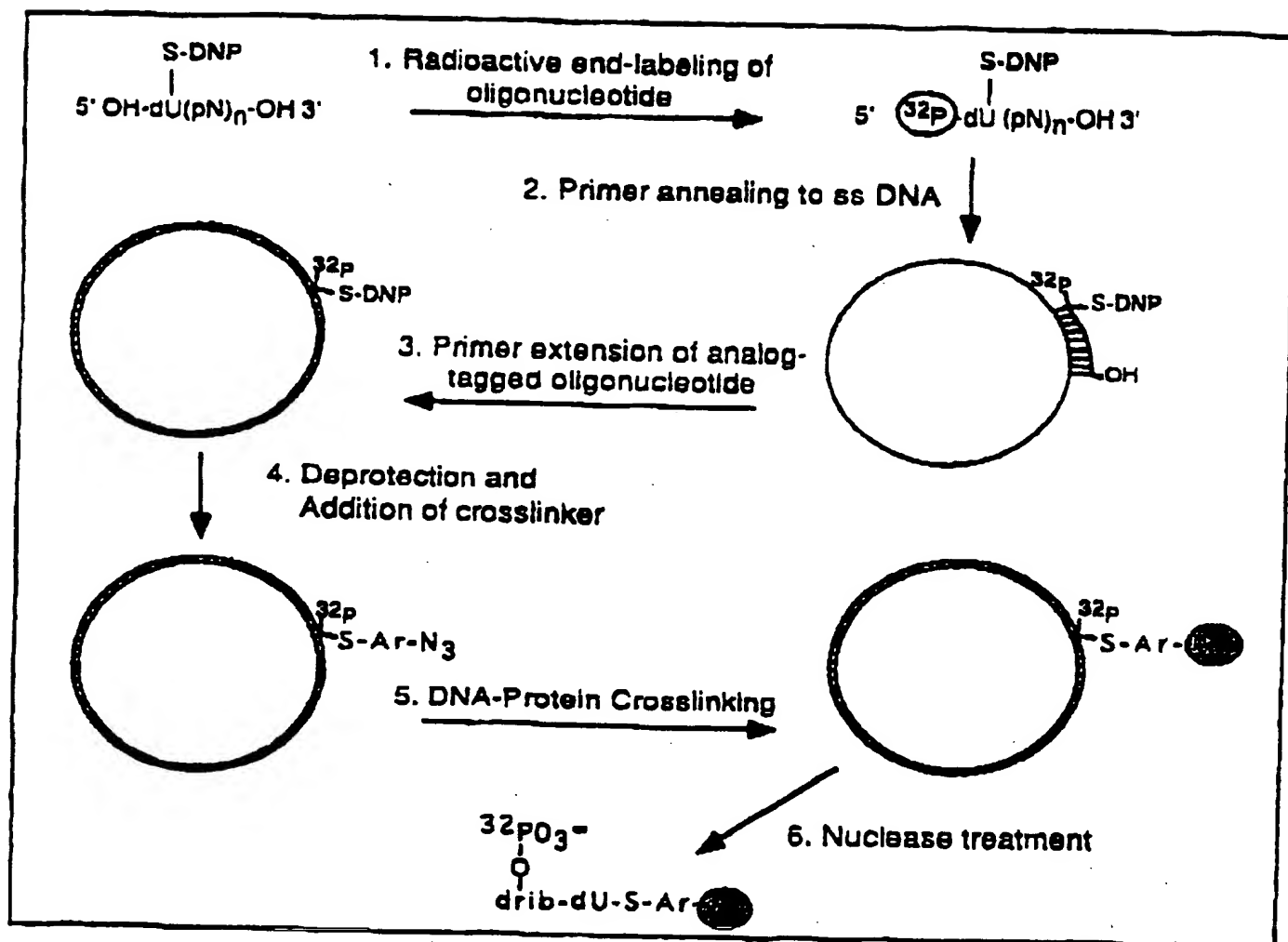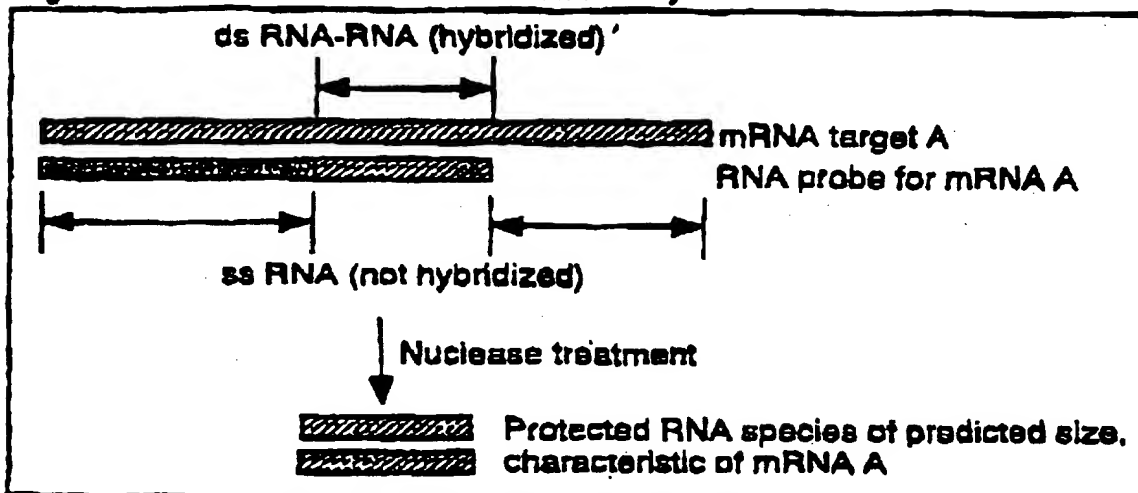
A.

B.

— XC

— BPB

← DNP-DNA

← Control DNA

6 / 7

Figure 6: Ribonuclease Protection Assay

Hatched areas indicate regions of RNA-RNA homology with the target gene (hybridization and double-stranded structure). Shaded areas in the RNA probe are not homologous to the target mRNA and therefore do not hybridize, leaving regions of single-stranded RNA. These ssRNA regions are susceptible to cleavage by certain nucleases.

FIGURE 7: DNA Sequence of Human Autocrine Motility Factor Receptor (AMFR)

```
   1 GAACTGCCT GTGAACATCT TTTCCACAAC TCCTGTCCTC GTTCCTGGCT AGAACAAGAC
  61 ACCTCCTGTC CAACATGCAG AATGTCTCTT AATATTGCCG ACAATATCGT GTCAGGGAAG
 121 AACATCAAGG AGAGAACTTG GATGAGAATT TTGGTTCCTGA TGCAGCAGCC GAAGGAGAACC
 181 TCGCCTTAAAC CAACACAATC ACTTCTTCCA TTTCGGATGGG TCTCGGATTG CGAGCCTGGCT
 241 GCCGGAGTTTT TCGGTTGAAG TGATGCACAC CACCAACATT CTTGGCCATTA CGCAGGGCCAG
 301 CAACTCCCAG CTCAATGCAA TGGCCTCATCA GATTCAAGAG ATGTTTCCCC AGGTTCCATA
 361 CCCATCTGGTA CTGCCAGGACC TCCAGCTGAC ACGCTCAGTT AAATAACAAC AGTCAATATT
 421 TTAGAAGGAC GGATTCAACT ACCTTTTCCT ACACAGCCGGT CAGATAGCAT CAGAACCTGCA
 481 TTGAACAGTC CTGTGGAAAG GCCAAGCAGT GACCAGGAAG AGGGAGAAAACTTCTGCTCAG
 541 ACCGAGCCGTG GTGCCACTCGG ACCTCAGTCC TCCCCTGGAG GAGACGCGTGG ACTTCGGCGGA
 601 GGTCGGAAGTG GAGCCCCAGTG AGGTGGAAGA CTTCGAGGCT CGTGGGAGCC GCTTCTCCAA
 661 GTCTCGTCTGAT GAGAGAACAGC GCATCGTCGT CGGAGGGTAA GGACTGAACTC CTCGAGCGAAG
 721 CTCGGCAAACG TTTCTTGAAC AAAAGTTCTG AAGATGAATCC GGCCCTCAGAG AGCTTCCTCC
 781 CCTCGGGAAGG TGCGTCCTCT GACCCGGTCGA CCCTGGGTCG AAGGATCGTG GCTGCCGGCGC
 841 GGAACCGGAGG CTTCAGAAGCA CGCAGGACCTC CTAGCGCTCC CCTTGCCTTCC TCAGCTGCCT
 901 CCTGCGGCCCT GTGCCCGGACT GACTGGGAGGA GGGCCTGTCCC AATTCTGCCG CTCCATGGAA
 961 AAGCCGGGCTT GACTGCAATTG CCGCTCGTATA AAGCATGTGG TCTTATAGTG TTTGGACAGC
1021 TGATAAATTTT AATCCTTCTTT TGTAATACTT TCTATGTGAC ATTTCTCTTCCCCTTAGAAA
1081 CACTGCAAAT TTTTAACTGTA GGTATGATCT CTTCTGCTTGT TGACTGGACT GCTTGGGCGTG
1141 GGGGACGGATC AGGAGGAAGT GAGCCAGTCG CCTGCCTGCA GCAGGCAGCCT TCTACTCCTTG
1201 CCTCATGCAT ACGTCCCACA AATGCCAGGTG TCCTGAGCAC CACACCCAGT GGGAAGAGTG
1261 TGGGGGAGGC GCACAGTGTG AGCCCGCGCC CACGTCGTGG GGTAACATCT GTTATCAAAC
1321 TGCTGTGGTT GTTGTGGAAG CATGTAGACT GTGCCAGAGC CAGACCCAGG GGCTCATGCA
1381 CCCCTGAGCA GCCAGGCCATC TTGGAAAAGG AACTCTTGGT TCGATACCTG GAGCAGAGGA
1441 GGGGAAAGTC CAGGGGCTATA GGGGTGTGATG AAGTCACCCC TTTCTGTCCC ACTACATGTG
1501 GGACTGACTT TCCGAGCCTC CAGTCCAAAG CCGGCTTGAT TTCCGTGAAC TCTGGTGCTC
1561 CTTGCATCTCA TGAGTGTGCC CCATGGGTCC CCTCCCCTCT CAGCCATTTCC TTGTCCCGTC
1621 TGGACCTGGG GAGTGGTTAG GCCAGCAAGCT TTGGCTTATG GTTTTCATTC ATTGGTGAAG
1681 TAAATTAGGC TAGTTGCACGT AAAGGCCGTG GGTTTGGTCC TTGAACAAGA TGTGGGGCTT
1741 GCCAAGATGGG AGAGTAAAGCCTTAAG
```

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6   C07H21/00   C07H19/10   C07H19/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6   C07H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 01445 A (RES CORP TECHNOLOGIES INC) 20 January 1994 see the whole document ---	1-4, 6-11, 13-18, 26,27, 29-36
X	M.M. HANNA: "Photochemical Cross-Linking Analysis of Protein-Nucleic Acid Interactions in Escherichia coli Transcription Complex from Lambda PR' Promoter" METHODS IN ENZYMOLOGY, vol. 274, 1996, pages 403-418, XP002045925 see the whole document ---	1-37

-/--

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

Date of the actual completion of the international search	Date of mailing of the international search report
6 November 1997	1 6. 12. 97

Name and mailing address of the ISA
        European Patent Office, P.B. 5818 Patentlaan 2
        NL - 2280 HV Rijswijk
        Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
        Fax: (+31-70) 340-3016
	Authorized officer
	
	Bardili, W

page 1 of 2

International Application No

PCT/US 97/12888

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	M. M. HANNA AND K.L. MEYER: "Synthesis and Characterization of a New 5-Thiol-Protected Deoxyuridine Phosphoramidite for Site-Specific Modification of DNA" BIOCONJUGATE CHEMISTRY, vol. 7, no. 4, 1996, pages 401-412, XP002045926 see the whole document ---	1-37
A	J. CHATTOPADHYAYA ET AL.: "Synthesis of C-5 and N-3 Arenesulfenyl Uridines" ACTA CHEM. SCAND. SER. B, vol. 39, 1985, pages 203-12, XP002045927 -----	

1

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9401445 A	20-01-94	NONE	

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: BASE-PROTECTED NUCLEOTIDE ANALOGS WITH PROTECTED THIOL GROUPS

(57) Abstract

The present invention is directed to protected thiol analogs of pyrimidine bases for syntheses of DNA and RNA by chemical or enzymatic methods. The subject analogs include reagents suitable for DNA or RNA synthesis via phosphoramidite, H-phosphonate or phosphotriester chemistry as well as reagents suitable for use by RNA and DNA polymerase, including thermostable polymerases employed by PCR or other nucleic acid amplification techniques. The nucleotide analogs synthesized by methods of this invention can thus be incorporated into oligonucleotides or polynucleotides, deprotected and derivatized with a functional group. In some cases the protecting groups are themselves antigenic and may be left on the oligonucleotides or polynucleotides for detection with antibodies. A method of synthesizing oligonucleotides with a functional group using the subject nucleotide analogs is also provided.

BASE-PROTECTED NUCLEOTIDE ANALOGS
WITH PROTECTED THIOL GROUPS

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT
Not Applicable

5 ## BACKGROUND OF THE INVENTION

The present application claims the benefit of the filing date of U.S. Provisional Application 60/022,573, filed July 24, 1996.

The present invention is directed to thiol-protected
10 pyrimidine nucleotide analogs which can be used, as one example, for syntheses of DNA and RNA by chemical or enzymatic methods. The subject analogs include reagents suitable for DNA or RNA synthesis via phosphoramidite, H-phosphonate or phosphotriester chemistry as well as reagents suitable for use by RNA and DNA
15 polymerases, including thermostable polymerases employed by PCR or other nucleic acid amplification techniques. Methods of synthesizing the nucleotide analogs are also provided by the present invention. The nucleotide analogs of this invention can thus be incorporated into oligonucleotides or polynucleotides,
20 deprotected, and then derivatized with a functional group.

Oligonucleotides with a variety of modifications have widespread utility for many purposes, such as stabilizing oligonucleotides to degradation, introducing reporter groups, allowing site-specific delivery of therapeutics, and introducing
25 crosslinkers. Such modifications can occur as modified internucleotide phosphate linkages or analogs of such linkages, modified sugars or modified bases. Additionally, 5'- or 3'-end conjugates of the oligonucleotides represent another class of modified oligonucleotides. The present invention relates to
30 base-modified nucleotide analogs with protected thiol groups; these analogs are intermediates for chemical or enzymatic synthesis of oligonucleotides and polynucleotides.

The synthesis of oligoribonucleotides and oligodeoxyribo-nucleotides (i.e. oligonucleotides) containing base-modified
35 nucleotides at specific positions provides a powerful tool in the analysis of protein-nucleic acid or nucleic acid-nucleic acid interactions. These oligonucleotides have many other

SUBSTITUTE SHEET (RULE 26)

potential uses, such as the site-directed delivery of therapeutics, utility as anti-sense therapeutics, and utility as diagnostic probes. Nucleotide analogs can be introduced into nucleic acids either enzymatically, utilizing DNA and RNA
5   polymerases, or chemically, utilizing manual or automated synthesis. Preparation of such oligonucleotides by automated synthesis utilizing, for example, phosphoramidite nucleotides allows for incorporation of a broad range of nucleotide analogs without the restraints for specific substrate conformation of
10   the nucleotides that is imposed by most polymerases. Often, nucleotide analogs containing photoreactive crosslinking groups are introduced into oligonucleotides to probe protein-nucleic acid interactions via photocrosslinking (for a partial review, see Hanna, 1989, Methods Enzymol. 180:383-409; and Hanna, 1996,
15   Methods Enzymol, 273 Chapter 31). Deoxyoligonucleo-tides containing 4-thiothymidine have been prepared and used for photochemical crosslinking of proteins directly to the nucleotide bases through the group (Nikiforov et al., 1992, Nucleic Acids Res. 20:1209-14). Similarly, oligonucleotides
20   containing 3-(amino-propyl)-2'-deoxyuridine have been prepared and the amino group subsequently modified with fluorescent, photoactive or other reporter groups (Gibson et al., 1987, Nucleic Acids Res. 15:6455-66). However, the former thiodeoxynucleotide suffers the disadvantage that if it is
25   modified with a thiol modifying reagent the normal Watson-Crick base of the nucleotide is drastically affected, making this and similar analogs generally unsuitable for use in enzymatic nucleic acid synthesis.

Other analogs involving modifications at the C5 position of
30   deoxyuridine have also been previously reported. The thiol-containing analog, 5-thiocyanatodeoxyuridine phosphoramidite, provided a 5-mercaptodeoxyuridine moiety within the oligonucleotide following reduction of the thiocyanate (Bradley & Hanna, 1992). However, the thiocyanato moiety displayed
35   variable stability during synthesis of both the nucleotide and the oligonucleotide and therefore this compound did not represent an ideal analog for incorporation of 5-thiol modified nucleotides into nucleic acids. The syntheses of a series of

3

phosphoramidites containing alkylthiol tethers at the C5 position of deoxyuridine has been reported (Goodwin & Glick, 1993). The thiol groups in these analogs are attached to the ring by either a three, four, or five carbon chain. The

5   presence of the carbon chains makes the minimal distance between the molecular probes and the oligonucleotide greater than that which can be achieved with our analog. In addition, these compounds represent alkyl thiol analogs which have a lower reactivity for modification of the thiol group than the 5-

10   mercaptopyrimidine analog. This is due to an increase in acidity of 5-mercaptopyrimidines (pKa~5-5.6) over alkylthiol moieties (pKa~8-10). Phosphoramidites containing alkylthiol tethers at the N3 position of thymidine have also been prepared, but the position of this modification results in a disruption of

15   the Watson-Crick base pairing. These analogs have been used mainly for preparing disulfide cross-links in DNA for studies involving stem loop and triple helical structures (Glick, 1991; Goodwin et al., 1994).

    Several 5-modified deoxyuridine phosphoramidites are

20   commercially available which contain functional groups (i.e. carboxylic acids or alkyl amines) for post-modification following incorporation into the oligonucleotide. Likewise, alkyl thiol ethers of 5-mercaptodeoxyuridine containing protected carboxylic acids and alkyl amines have also been

25   described (Bergstrom et al., 1991). The protected functional groups in these analogs are not attached directly to the ring but are positioned at the end of carbon chains. These groups are not as easily modified as mercaptans: the functional groups formed during post-synthetic modification are limited, and an

30   easily cleavable group is not available. Thiol-containing phosphoramidites for incorporating 4-thiothymidine (Clivio et al., 1992; Xu et al., 1992c), 4-thiodeoxyuridine (Clivio et al., 1992; Coleman & Kesicki, 1994), 2-thiothymidine (Connolly & Newman, 1989; Kuimelis & Nambiar, 1994), 6-thiodeoxyguanosine

35   (Christopherson & Broom, 1991; Waters & Connolly, 1992; Xu et al., 1992b) and 6-thioinosine (Clivio et al., 1992a) have been reported. These analogs can occupy internal positions within an oligonucleotide and can serve as photochemical crosslinkers.

SUBSTITUTE SHEET (RULE 26)

4

However, they cannot be further modified without disrupting Watson-Crick base pairing, and therefore, as photocrosslinking probes, these analogs are only useful for evaluating interactions which occur directly with the nucleotide base. Modification with other molecular probes (e.g., fluorescent tags) would also disrupt Watson-Crick base pairing. In addition, the deprotection of oligonucleotides containing these thiol-modified nucleotides must be carefully monitored to prevent conversion of these analogs to the corresponding oxygen and nitrogen derivatives.

Described herein are nucleotide analogs which can be used for site-specific modification of DNA or RNA at internal and terminal positions within the DNA or RNA sequence, and after modification, for molecular probes which can be placed at variable distances from the DNA or RNA backbone.

The present invention provides novel base-protected nucleotide analogs, both ribonucleotides and deoxynucleotides, that contain masked thiol groups on the 5 position of pyrimidines, which is not involved in Watson-Crick base pairing. These analogs can be incorporated into oligonucleotides via automated synthesis and isolated with the thiol protecting group intact. After removal of the thiol protecting group many types of functional groups, such as photocrosslinking agents, fluorescent tags, radioisotopes, biotin, reporter molecules, spin labels (e.g., commercially available proxyl or tempo), chemiluminescent, antigenic or other functional groups, can be site-specifically attached by utilizing thiol-modifying reagents. This feature adds a level of specificity to the oligonucleotide modifications not present with the amino-tagged analogs previously described (Gibson et al, 1987), and enables examination of molecular interactions that are not directly at the nucleotide base by allowing functional groups to be placed at varying distances from the base or helix strand. Since these analogs have the functional group attached via the sulfur atom, some have the further advantage of being cleavable under conditions which will not degrade or modify the oligonucleotide.

SUMMARY OF THE INVENTION

This invention relates to pyrimidine nucleotide analogs which contain modified bases with protected thiol groups attached at a position on the base, preferably the 5 position, which is not involved in Watson-Crick base pairing. These nucleotide analogs are intermediates in chemical or enzymatic synthesis of DNA or RNA oligonucleotides and are therefore stable under conditions required for synthesis of these molecules. After synthesis, the protecting group on the analog is removable to generate a reactive thiol group. Once generated, the thiol group can be treated with thiol modifying reagents to attach functional groups such as crosslinking agents or reporter molecules.

In particular, the nucleotide analogs of the present invention have the formula:

wherein

R₁ is -H, -OH, a mono, di, or triphosphate group, or -OR₄;

R₂ is -H, -OH, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, -OR₄, or a solid support bonded via an O at the 3' position;

R₃ is -H, -OH, a mono, di, or triphosphate group, or -OR₄;

R₄ is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at the 5 position on said base that is not involved in Watson-Crick base pairing or does not disrupt normal Watson-Crick base pairing, said protected thiol group being stable under conditions of chemical nucleic acid synthesis and/or conditions of enzymatic nucleic synthesis and being convertible to a reactive

6

thiol (SH) after said synthesis.

Preferred nucleotide analogs of the present invention are the protected phosphoramidites or 5' mono, di and triphosphates of modified cytosine or uridine bases for use in the chemical synthesis of DNA and RNA by the phosphoramidite method or enzymatic synthesis with polymerases, ligases, or other nucleotide polymerizing enzymes.

In addition, the nucleotide analogs of the present invention include other nucleoside phosphates, containing 3', 5', or 3',5' monophosphates, diphosphates, or triphosphates and further optionally comprising 2' mono, di, or triphosphates when the nucleoside is a ribonucleoside.

Another aspect of this invention provides nucleic acids and oligonucleotides containing the subject nucleotide analogs having a protected thiol group on a base moiety of that nucleic acid or oligonucleotide. A method is also provided to synthesize these nucleic acids or oligonucleotides, deprotect the thiol group and attach a functional group to the reactive thiol moiety.

Yet another aspect of this invention is directed to a method of synthesizing the subject base-protected nucleotide analogs.

DESCRIPTION OF THE DRAWINGS

Figure 1: Synthetic pathway for the DNP-labeled phosphoramidite (compound 7). Reaction conditions are as follows a) SCNCl,CH$_3$CO$_2$H; b) DTT,MeOH,EDTA; c) 2,4-dinitrofluorobenzene,Et$_3$N,CH$_3$CN; d) NaOMe,MeOH; e) DMT$^+$BF$_4^-$,DBMP,CH$_3$CN; f) 2-cyanoethyl-N,N,N',N'-tetraisopropyl-phosphorodiamidite,tetrazole,CH$_3$CN.

Figure 2: A. RP HPLC chromatograms of Oligonucleotides A and B. The purified 7-mers were analyzed by analytic HPLC on an Ultrasphere ODS column, eluted with a 15 minute linear gradient from 5% to 20% (v/v) acetonitrile in 20 mM TEAA (pH 7.1). (a) Oligonucleotide A, GTA TGT A, eluted at 9.8 min. (b) DNP-labeled oligonucleotide B, GTA T*GT A, eluted at 11.6 min, where T represents incorporation of the DNP-analog described in Figure 1. B. RP HPLC profile of enzymatic degradation of

SUBSTITUTE SHEET (RULE 26)

7

Oligonucleotides A and B.   The 7-mers were subjected to
enzymatic treatment with Nuclease P1 and CIP.   The reaction
mixture was analyzed by analytical HPLC on an Ultrasphere ODS
column, eluted with a linear gradient of acetonitrile in 100 mM
5   TEAA (pH 7.1) from 0% to 5% (v/v) in 5 min, 5% to 6% (v/v) in 15
min., 6% to 50% (v/v) in 5 min and finally 50% to 80% (v/v) in
15 min.   (a) oligonucleotide A reaction.   (b) DNP-labeled
Oligonucleotide B.

       Figure 3:   A.: Gel electrophoretic analysis of DNP-labeled
10   oligonucleotides.   The [^{32}P] labeled oligonucleotides C-F (Table
1) were analyzed by electrophoresis on a 25% polyacrylamide-7M
urea gel (acrylamide/methylene bisacrylamide=19:1).   Lane 1:
Commercial 24-mer marker, CGC GGA ATT CTC ATG CAT TGC CCA (SEQ.
ID. NO:5);   Lane 2; 24-mer, Oligonucleotide E;   Lane 3: DNP-
15   labeled 24-mer, Oligonucleotide F;   Lane 4: 22-mer,
Oligonucleotide C;   Lane 5: DNP-labeled 22-mer, Oligonucleotide
D.   B.: Gel electrophoretic analysis of the deprotection and
alkylation reactions.   Oligonucleotides E and F were reacted
with 1.4 M BME.   Half of these reactions were alkylated with 5-
20   IAF.   The treated oligonucleotides were analyzed by
electrophoresis on a 25% polyacrylamide-7M urea gel (acrylamide/
methylene bisacrylamide=19.1).   Lane 1: Oligonucleotide E + 1.4
M BME at room temperature;   Lane 2: Lane 1 reaction + 5-IAF;
Lane 3: Oligonucleotide E + 1.4 M BME at 45°C;   Lane 4: Lane 3
25   reaction + 5-IAF;   Lane 5: Oligonucleotide F + 1.4 M BME at room
temperature;   Lane 6: Lane 5 reaction + 5-IAF;   Lane 7:
Oligonucleotide F + 1.4 M BME at 45°C;   Lane 8: Lane 7 reaction
+ 5-IAF.

       Figure 4:   Analysis of PCR products.
30   A. PCR products were analyzed by electrophoresis on a 1.5%
agarose gel.   DNA was detected with ethidium bromide.   Lane 1:λ
Hind III markers;   Lane 2: PCR product using Oligonucleotide E
as a primer;   Lane 3: PCR product using DNP-labeled
Oligonucleotide F as a primer.   XC represents xylene cyanol and
35   BPB represents bromophenol blue.
       B. Dot Blot immunodetection of DNP-labeled DNA.   PCR
products were attached to a nylon membrane, incubated with a
primary antibody to DNP and a 2° antibody linked to horseradish

SUBSTITUTE SHEET (RULE 26)

peroxidase. Chemiluminescence was used for detection. Upper panel: PCR product using DNP-labeled Oligonucleotide F as a primer; Lower panel: PCR product using Oligonucleotide E as a primer.

5   **Figure 5:**   Protein Crosslinking to Site-Specifically Modified DNA

To analyze the DNA-protein complexes in the transcription complex a deoxynucleoside phosphoramidite was incorporated site-specifically into a single-stranded DNA oligonucleotide via
10   automated synthesis. In these studies, the nucleotide analog was modified with a masked reactive thiol group on the non-basepairing 5 position of deoxyuridine (5-S-DNP-dU, Figure 1). The oligonucleotide was then radioactively labeled uniquely on the 5' phosphate of the nucleotide analog with polynucleotide
15   kinase and [γ-^{32}P] ATP (Step 1). The radioactively labeled, base-modified oligonucleotide was then annealed to a single-stranded DNA template (Step 2) and incorporated site-specifically into double stranded DNA by primer extension (Step 3). The protecting group was then removed from the double-
20   stranded DNA by treatment with β-mercaptoethanol (Step 4), and after removal of the BME, a functional group was added to the unmasked thiol group. For crosslinking, the DNA was reacted with azidophenacyl bromide to attach a photoreactive azide group (Step 4). RNA polymerase and transcription factors are then
25   added to assemble a complex on the modified DNA, and the complex is irradiated with ultraviolet light to covalently join the DNA to proteins which are bound adjacent to the crosslinker (Step 5). Lastly, the DNA-protein complex is treated with a nuclease which leaves 5' NMPs to digest the DNA to a small fragment (Step
30   6). This will leave a single radioactive label on the 5' phosphate of the nucleotide involved in the DNA-protein crosslink.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a series of novel
35   nucleotide analogs which are masked synthons for use as intermediates in chemical or enzymatic synthesis of nucleic acids, including synthesis of both oligonucleotides and

9

polynucleotides. The nucleotide analogs of this invention, which contain a protected thiol group, can thus be incorporated into DNA or RNA under standard synthetic conditions without loss of the thiol protecting group. This stability of the thiol

5   protecting group permits site-selective introduction into a nucleic acid of the nucleotide analog in a manner which facilitates later addition of a functional group at that site. Thus, the subject oligonucleotides (or polynucleo-tides) can contain one (or more) of the subject nucleotide analogs.

10   Thus, functional groups which can not withstand the conditions for chemical nucleic acid synthesis, especially during automated synthesis, or which may be too bulky or sterically hindered for enzymatic nucleic acid synthesis can be readily incorporated into the final oligonucleotide (or

15   polynucleotide) product. Such derivatized nucleic acids have great utility in studying protein-nucleic acid or nucleic acid-nucleic acid interactions, as well as the potential for diagnostics and delivery of therapeutics. Hence, the present invention permits the skilled artisan to place chemical tags

20   such as crosslinking groups, fluorescent molecules, radioisotopes or other reporter molecules at specific positions on nucleic acid molecules for analysis of molecular mechanisms, for creation of diagnostic probes, for therapeutics, for antisense therapeutics and for many other purposes.

25   In particular, the nucleotide analogs of the present invention are intermediates for the chemical synthesis of DNA and RNA by manual or automated techniques and are represented by the formula:

wherein

30   R_1 is -H, -OH, a mono, di, or triphosphate group, or -OR_4;

        R_2 is -H, -OH, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite

10

group, a phosphonate group, an O-substituted
monophosphate group, $-OR_4$, or a solid support
bonded via an O at the 3' position;

R₃ is -H, -OH, a mono, di, or triphosphate group, or
$-OR_4$;

R₄ is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected
thiol group attached at a position on said base,
preferably the 5 position, that is not involved
in Watson-Crick base pairing or does not disrupt
normal Watson-Crick base pairing, said protected
thiol group being stable under conditions of
chemical nucleic acid synthesis and/or conditions
of enzymatic nucleic synthesis and being
convertible to a reactive thiol after said
synthesis.

As used herein, protecting groups defined by R₄ are known
to those of ordinary skill in the art and include any known
protecting group suitable for protection of the 2', 3' or 5'
hydroxyls of ribose and the 3' or 5' hydroxyls of deoxyribose
sugars. In this regard, Greene et al. (1990) Protecting Groups
in Organic Synthesis, 2nd ed., John Wiley & Sons, Inc., New
York, NY, provides a comprehensive review of protecting groups,
and methods of preparing the corresponding protected compounds,
which can be used for different reactive groups, including
reactive hydroxyl groups.

Accordingly, examples of protecting groups defined by R₁,
R₂ and R₄ are lower alkyl, lower cyanoalkyl, lower alkanoyl,
aroyl, aryloxy, aryloxy-lower alkanoyl, haloaryl,
fluorenylmethyloxy-carbonyl (FMOC), levuloyl, 9-phenylxanthene-
9-yl, trityl, p-monomethoxytrityl (MMTr), p-dimethoxytrityl
(DMTr), isopropyl, isobutyl, 2-cyanoethyl, acetyl, benzoyl,
phenoxyacotyl, halophenyl, 1-(2-chloro-4-methylphenyl)-4-
methoxy-4-piperidinyl, 2'-acetal, O-nitrobenzyl, tert-
butyldimethylsilyl (TBDMS), tetrahydrofuranyl, 4-
methoxytetrahydropyranyl and related groups.

Preferred R₁ protecting groups include 5'OH protecting
groups especially DMTr, MMTr, FMOC, levuloyl and 9-

phenylxanthene-9-yl groups, and most preferably DMTr. Preferred
R_2 protecting groups include 3'OH protecting groups, especially
the acetyl group. Preferred R_4 protecting groups include 2'OH
protecting groups, especially 2'-acetal, tetrahydrofuranyl, 4-

5   methoxytetrahydropyranyl and 1-(2-chloro-4-methylphenyl)-4-
methoxy-4-piperdinyl groups, and most preferably TBDMS.

As defined herein, solid supports include controlled pore
glass (CPG), polystyrene silica, cellulose, nylon, and the like.
Preferred solid supports are CPG and polystyrene. An especially

10 preferred solid support is CPG.

R_2 can be taken with the oxygen atom to which it is
attached, to form a phosphoramidite, phosphorothioamidite, a
phosphonate, an O-substituted monophosphate, or any other group
compatible with chemical nucleic acid synthesis, especially

15 automated DNA or RNA synthesis. As used herein, the
phosphoramidite and phosphorothioamidite groups have the general
formulas I and II, respectively:

wherein R_6 is lower alkyl, n-cyano alkyl, substituted lower

20 alkyl, aryl, aralkyl, substituted aralkyl and the like, and R_7
and R_8 are independently lower alkyl or when taken together with
the nitrogen to which they are attached comprise cyclic groups
including:

25 preferably the R_6 group of the phosphoramidite group (Formula I)
is 2-cyanoethyl (CED) or methyl. Accordingly, the preferred
phosphoramidites of this invention are also referred to as CED
phosphoramidites or methyl phosphoramidites.

SUBSTITUTE SHEET (RULE 26)

12

As used herein, the term lower alkyl, when used singly or in combination, refers to alkyl groups containing one to six carbon atoms. Lower alkyl groups can be straight chain or branched and include such groups as methyl, ethyl, propyl, isopropyl, n-butyl, sec-butyl, isobutyl, t-butyl, pentyl, isopentyl, neopentyl, hexyl and the like. The preferred alkyl groups contain one to four carbon atoms.

The term aryl, when used alone or in combination, refers to an aromatic ring containing six to ten ring carbon atoms. The aryl group includes phenyl, and 1- or 2-naphthyl. The preferred aryl group is phenyl.

The term aralkyl refers to aryl groups as described above to which substituents are attached to the aryl by an alkylene bridge. The most preferred aralkyl group is benzyl.

When R_7 and R_8 are lower alkyl, the preferred groups are each isopropyl groups.

As used herein a phosphonate group (or H-phosphonate) is represented by the formula:

$$O=\overset{\overset{\displaystyle |}{O}}{\underset{\underset{\displaystyle O^-}{|}}{P}}-H$$

and may also be provided as salts, and preferably as triethylammonium salts.

As used herein, O-substituted monophosphates have the formula:

$$R_9-O-\overset{\overset{\displaystyle |}{O}}{\underset{\underset{\displaystyle O^-}{|}}{P}}=O$$

wherein R_9 is lower alkyl, haloalkyl, aryl, haloaryl, or heteroaromatic. By haloalkyl or haloaryl is meant alkyl or aryl groups, respectively, which have been substituted with one or more halogen atoms including F, Cl, Br, or I. Preferred halo

SUBSTITUTE SHEET (RULE 26)

substituents are Cl and Br. Preferred R_4 groups include 2-chlorophenyl, 2,5-dichlorophenyl, 2,2,2-trichloroethyl and the N-oxide of 4-methoxypyridine-2-methylene groups.

Modified pyrimidine bases have a protected thiol group attached at a position on the base which is not involved with Watson-Crick base pairing or which does not disrupt normal Watson-Crick base pairing. The protected thiol groups of this invention are stable, i.e. not removable, under the conditions used for chemical synthesis of DNA or RNA, and particularly, under the conditions employed in automated DNA or RNA synthesis. Furthermore, the protecting group of the thiol is removable under conditions which do not disrupt the integrity of the oligonucleo-tide or polynucleotide. In other words, after a nucleotide analog of the present invention has been incorporated into an oligonucleotide, for example, the protected thiol group can be converted to a reactive thiol (SH) to which functional groups can subsequently be added using thiol-modifying reagents.

In accordance with this invention the protected thiol groups include, but are not limited to, the dinitrophenyl group. As used herein, Watson-Crick base pairing refers to the hydrogen bonding pattern of adenine-thymine (AT) base pairs, adenine-uracil (AU) base pairs, or of guanine-cytosine (GC) base pairs. Accordingly, the preferred protected thiol position on pyrimidines is the 5 position on the pyrimidine (C,U) ring.

The preferred bases of the present invention are thus pyrimidines such as cytosine (C), and uracil (U). Cytosine bases contain an additional reactive group, specifically an exocyclic amine, which must also be protected during assembly of the nucleotide chain via chemical synthesis. Accordingly, bases of this invention can have additional protecting groups attached as needed to any of the ring positions. Such protected bases are well known in the art and include but are not limited to, N^2, N^4-anisoyl cytosine, N^4-benzoyl cytosine, and N^4-isobutyryl cytosine. The N^4-isobutyryl cytosine is known by the trade name FOD^{TM} base protection (Applied Biosystems, Inc.). Use of these further protected bases is compatible with incorporation of the protected thiol groups as well as subsequent deprotection reactions.

14

The bases of this invention also include any related base analog that is capable of base pairing with a guanine or adenine, the corresponding protected analogs as set forth above for use in chemical synthetic methods to produce DNA and RNA. For example, such base analogs include, but are not limited to, pseudocytosine, isopseudocytosine, 4-acetylcytosine, 2'-O-methylcytosine, dihydro-uracil, 2'-O-methyluracil, 2'-O-methyl-pseudouracil, 1-methyl-pseudouracil, 3-methylcytosine. Bases attached to a ribose or deoxyribose sugar in an α-anomeric configuration can also be present.

In a preferred embodiment the bases of the invention are C, U, and N^4 protected C.

The compounds of the present invention also include nucleoside phosphates of the formula:

wherein:

R_1 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R_2 is -H, -OH, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, -OR$_4$, or a solid support bonded via an O at the 3' position;

R_3 is -H, -OH, a mono, di, or triphosphate group, or -OR$_4$;

R_4 is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at the 5 position on said base.

In the case of the modified bases the protected thiol group is also stable, i.e. not removable, under the conditions used for enzymatic synthesis of DNA or RNA. Additionally, the pyrimidine bases generally do not require protection of the

additional reactive groups such as the exocyclic amines for enzymatic incorporation into DNA or RNA. Accordingly, the preferred bases for this class of compounds includes C and U.

In accordance with this invention the phosphate groups of R_1, R_2 and R_3 embody all the phosphorylated forms at the C_3' and C_5' positions of the sugar moiety (and C_2' position when the sugar is ribose) and include monophosphates, diphosphates, triphosphates and tetraphosphates. Preferably the phosphate group is a triphosphate for R_1.

The preferred compounds of this invention include the phosphoramidites and 5'triphosphates of 5-SDNP-dU, 5-SDNP-U, 5-DNP-dC, and 5-SDNP-C. In the case of the phosphoramidites, the base moieties can contain additional protecting groups on the exocyclic amines. The preferred compounds also include the 5'triphosphates of 5-S-S-Et-dU, 5-S-S-Et-U, 5-S-S-Et-dC and 5-S-S-Et-C, 5-S-CH$_2$-NHCOCH$_2$Ph-dU, 5-S-CH$_2$-NHCOCH$_2$Ph-U, 5-S-CH$_2$-NHCOCH$_2$Ph-C, 5-S-CH$_2$-NHCOCH$_2$Ph-dC, 5-SCH$_2$CH$_2$DNP-dU, 5-SCH$_2$CH$_2$DNP-U, 5-SCH$_2$CH$_2$DNP-dC, and 5-SCH$_2$CH$_2$DNP-C.

The nucleotide analogs of the present invention can be prepared by adding a protected thiol group to the base moiety of the desired nucleoside. The so-modified nucleoside can then be phosphorylated to produce a nucleotide analog phosphate compound of this invention using conventional phosphorylation techniques. To produce a phosphoramidite, phosphonate, phosphorothioamidite or O-substituted monophosphate of this invention, the 5'-OH or 2'-OH groups of above-modified nucleoside are protected as necessary by addition of the desired protecting group by standard methodology. This protection step(s) is (are) followed by conversion to the nucleotide-analog phosphoramidite, phosphorothioamidite, phosphonate or O-substituted monophosphate by reaction of the 3'OH of the nucleoside with the appropriate modifying group. Similarly, the 3'OH can be attached to a solid support, such as CPG, or another protecting group using conventional methodology available to the ordinarily skilled artisan in this field.

For example, to prepare protected thiol-groups at the 5 position of uridine or 2'-deoxyuridine nucleosides, the sulfur atom must be incorporated on the base, for which the method of

16

Nagamachi et al. (1972, J.C.S. Chem. Commun. 18:1025-6) has been modified. An equivalent amount of potassium thiocyanate is dissolved into a solution of chlorine gas in glacial acetic acid to produce the thiocyanogen chloride in solution. An iodometric

5   titration is performed with Cl_2/HOAc to determine the amount of Cl_2 present. The solution can be used directly or filtered prior to the next step. In either case, the acetyl-protected U or dU is added all at once with stirring. This reaction is maintained with stirring at room temperature for about 1.5 hours until the

10   reaction is completed. An excess of cyclohexene or other quenching agent is then added to quench any remaining thiocyanogen chloride. Quenching is complete in 15 to 60 min, and usually in about 30 min. After removing solvents and organic residue in vacuo, the remaining residue can be purified

15   by silica gel chromatography with a gradient of petroleum ether and ethyl acetate or with other chromatographic methods. These acetyl-protected 5-thiocyanato dU and U nucleosides can then be reduced to 5-mercapto nucleosides for further derivatization as described herein to produce the nucleotide analogs of the

20   present invention.

The mono-, di-, tri or tetra phosphate compounds can be prepared by a similar reaction by beginning with the appropriate starting 5-thioprotected nucleoside. If necessary, the 2'-OH of the ribose sugar can be protected prior to the reaction using

25   any of the known 2'-OH protecting groups by conventional techniques.

The described synthesis and utilization of 5-thiocyanato-deoxyuridine phosphoramidite for placement of the sulfur directly on the base has been previously described (Bradley &

30   Hanna, 1992). However, the synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-S-(2,4-dinitrophenyl) mercapto-2'-deoxyuridine-3'-O-(2-cyanoethyl-N,N'-diisopropyl) phosphoramidite is more reproducible and the product is more stable than 5-thiocyanatodeoxyuridine phosphoramidite and

35   represents a more desirable analog. A variety of functional groups can be attached to the reactive thiol. Once this modified nucleotide is incorporated into an oligonucleotide or nucleic acid, the protecting group can be removed to unmask the

17

reactive thiol, making it available for further derivitization with thiol-specific agents. This method has the advantage that an oligonucleotide tagged with this analog can be synthesized and stored for long periods without removal of the protecting group. When needed, an aliquot can then be deprotected and modified. In addition, in some cases the protecting groups are themselves antigenic (e.g., ⁻DNP) and may be retained in the nucleic acid product for detection utilizing antibodies.

One method used to produce one of the nucleotide analog phosphoramidites of this invention and described in detail in Example 1 below is shown in Fig. 1. Briefly, as depicted, 5-SDNP-dU or 5-SDNP-U are prepared by reduction of the SCN, protection of the thiol as the DNP ether, and de-esterification of the sugar ring. The protected-thiol nucleoside thus formed is then reacted with the DMT⁺BF₄⁻, or another 5'OH protecting group, in an anhydrous organic solvent in the presence of an organic base for 2-24 hours until the reaction is complete. The resulting product can be isolated by liquid chromatographic methods and then converted to a CED phosphoramidite by reaction with 2-cyanoethyl N, N, N', N'-tetraisopropyl phosphordiamidite and tetrazole under anhydrous conditions and in an inert atmosphere. This reaction is preferably stirred for 1 hour, although this time can be varied, the solvents removed in vacuo and the residue purified by chromatographic methods. Exposure to atmosphere is acceptable but should be minimized. This phosphoramidite can be stored dry under positive pressure of inert atmosphere (argon or nitrogen) in a tightly sealed container at low temperature.

Other phosphoramidites of this invention can be prepared by reaction with the appropriate chlorophosphoroamidite. For example, the O-methylphosphoramidites can be prepared by reacting the 5-SDNP-5'-DMTr-dU with N,N-diisopropylmethylphosphonamidic chloride by conventional techniques. Similarly, the phosphorothioamidite, phosphonates, O-substituted monophosphate can be prepared using conventional techniques with commercially available reagents. As mentioned above to avoid unwanted side reactions, protection of the 5'OH and 2'OH groups on the sugar moiety and the exocyclic amine

SUBSTITUTE SHEET (RULE 26)

groups on the base moiety may be necessary before the final reaction step which produces the nucleotide analogs of this invention.

In general the chemical synthetic routes to nucleotide phosphoramidites, phosphorothioamidites, phosphonates and O-substituted monophosphates, as well as nucleotide phosphates, are well known. In addition, nucleoside phosphates can be enzymatically synthesized. Chemical synthetic techniques for these compounds, as well as the common synthetic routes to prepare RNA and DNA, have been described in many sources. Particularly, useful references include Gait (1984) Oligonucleotide Synthesis: A Practical Approach, IRL Press, Oxford; Blackburn et al. (1990) Nucleic Acids in Chemistry and Biology, IRL Press, Oxford, especially Chap. 3; Chaps 13-16 of Methods in Enzymology, Vol. 154 (Wu et al., eds.) Academic Press, San Diego, CA, 1987; Chaps. 13-14 in Methods in Molecular Biology, Vol. 4 (Walker, ed.), Humana Press, Clifton, NJ, 1988; and Uhlmann et al. (1990) Chem. Rev. 90:544-584.

In addition to providing methods for chemical synthesis of DNA and RNA, some of these references (particularly Gait and Uhlmann et al.) describe the reactions and methodology for adding protecting groups to 5'OH, 3'OH and 2'OH groups, and exocyclic amine groups (N^6 of adenine, N^4 of cytosine, N^2 of guanine). These references also provide information and protocols to attach nucleotides to solid supports which protocols are useful for attaching the base-protected nucleotide analogs of this invention.

The synthesis of 5-SDNP-dC and 5-SDNP-C can be accomplished by reaction of 5-Br-dC or 5-Br-C with sodium hydrogen sulfide, as described by Solan, followed by conversion into the DNP thiol ether. In these cases, it may be necessary to first protect the 4 amino group of cytosine (e.g., as the iso-butoxy amide) and complete the remaining protection steps (5'-O-DMTr and 2'-O-TBDMS, if needed) before preparation of a CED phosphoramidite or conversion to the triphosphate (Sung, 1982, J. Org. Chem. 47:3623). Alternatively, 5-S-dU or 5-S-U can be converted to the corresponding compounds. Several methods are available for this conversion of U to C, including Sung, Xu, and MacMillan.

SUBSTITUTE SHEET (RULE 26)

19

The conversion of 5-SDNP-dU and 5-SDNP-U to 5'phosphates,
phosphoramidites, phosphorothioamidites, phosphonates, and O-
substituted monophosphates has been described herein above. All
of these reaction schemes can be used to produce the
corresponding thiol protected cytosine analogs. If necessary,
various protecting groups for the 5'OH, 3'OH or 2'OH groups as
well as the exocyclic amines can be added in accordance with the
methodology described herein.

Another aspect of this invention relates to the
oligonucleotides or polynucleotides containing the nucleotide
analogs of this invention and a method of preparing such nucleic
acids using the subject nucleotide analogs. Oligonucleotides
and polynucleotides of this invention are made by standard
methods of chemical (automated or manual) synthesis or enzymatic
synthesis of DNA and RNA. Such methods are well known in the
art. In chemical synthesis, the nucleotide analog of this
invention is substituted for a particular nucleotide at the
desired point in the synthesis.

After incorporation of the nucleotide analog and complete
synthesis of the oligonucleotide or polynucleotide, the thiol
can be deprotected and reacted with any number of thiol-
modifying reagents to attach a functional group at that point on
the oligonucleotide or polynucleotide. Deprotection of S-DNP or
S-S-Et can be accomplished by treatment with β-mercaptoethanol
or by other means of reducing sulfides. In a preferred method
for the DNP analog, deprotection is accomplished by treating the
oligonucleotide with 1.4M BME at 45°C for 4 hours or at room
temperature overnight.

Any variety of functional groups can then be added to the
reactive thiols group generated by the deprotection step. Such
functional groups include cross-linking groups, photoactive
cross-linking groups (e.g. arylazides), and reporter molecules
such as radioisotopes, biotin, enzymes and fluorescent markers.
The methods for adding functional groups are well know in the
art.

Yet another aspect of this invention provides a method of
preparing the nucleotide analogs of the invention. In
particular, this method involves preparing a thiol protected

nucleoside or nucleotide base wherein said thiol is attached to a position on said base that is not involved in Watson-Crick base pairing or does not disrupt normal Watson-Crick base pairing; reacting said nucleoside or nucleotide base under

5   conditions to effect conversion of said base to a phosphoramidite, phosphorothioamidite, phosphonate, o-substituted monophosphate or phosphate nucleotide analog and under conditions which do not destroy the protected thiol; and recovering said analog. In accordance with this invention, this

10   method is accomplished as described above for synthesis of the subject nucleotide analogs. Recovery of the analogs can be accomplished by HPLC, FPLC or other chromatographic separation techniques.

     A further aspect of this invention provides a method of

15   synthesizing a nucleic acid with a functional group by incorporating a thiol protected nucleotide analog in accordance with this invention into a nucleic acid by a chemical or enzymatic method for nucleic acid synthesis; recovering the nucleic acid containing the analog; deprotecting the analog of

20   that nucleic acid to produce a nucleic acid containing a reactive thiol group; reacting the reactive thiol group with a thiol-modifying reagent to thereby attach a functional group and produce the nucleic acid with the functional group; and recovering the nucleic acid with the functional group. In

25   accordance with this invention, this method is accomplished as described hereinabove for synthesis of oligonucleotides and polynucleotides (See for example Gait or Sambrook et al). As used herein, nucleic acids include oligonucleotides and polynucleotides. Preferably, the oligo-nucleotides range in

30   size from about 5 to about 100 nucleotides. Polynucleotides range in size from 100 nucleotides to 10 kb or more. Recovery of the analogs can be accomplished by HPLC, FPLC, other chromatographic techniques, extraction, phase separation or precipitation.

35   While the invention will now be described in connection with certain preferred embodiments in the following examples so that aspects thereof may be more fully understood and appreciated, it is not intended to limit the invention to these

21

particular embodiments.   On the contrary, it is intended to cover all alternatives, modifications and equivalents as may be included within the scope of the invention as defined by the appended claims. Thus the following examples, which include

5   preferred embodiments, will serve to illustrate the practice of this invention, it being understood that the particulars shown are by way of example and for purposes of illustrative discussion of preferred embodiments of the present invention only and are presented in the cause of providing what is

10 believed to be the most useful and readily understood description of formulation procedures as well as of the principles and conceptual aspects of the invention.

EXAMPLES

Procedures

15   **Materials.** Unless otherwise stated, starting materials for the chemical synthesis of the phosphoramidite were obtained from Aldrich, Sigma, Fluka or Fisher Scientific and were used without further purification. Anhydrous HOAc was prepared by fractional freezing (Nagamachi et al., 1974).   Anhydrous CH_2Cl_2 was

20 distilled from phosphorus pentoxide.   Anhydrous CH_3CN was purchased from Cruachem.   Dowex 50W x4 was purchased from BioRad.   Analytical thin layer chromatography was performed on Whatman silica gel 60 plates with fluorescent indicator. Column chromatography was performed with silica gel 60A from American

25 Scientific Products. Reagents for the oligonucleotide syntheses were obtained from Cruachem.   Standard exocyclic amine protecting groups were used: Benzoyl for dA and dC and isobutyryl for dG.   Solvents utilized for the reverse phase chromatographic purification of the oligonucleotides were of

30 HPLC grade.   The fluorescent thiol modifying agent, 5-IAF was purchased from Molecular Probes (Eugene, Oregon).

        Buffers. The buffers used were as follows: buffer A, 100 mM TEAA (pH 7.1); buffer B, 20 mM TEAA (pH 7.1); buffer C, 50 mM Tris-HCl (pH 4.9), 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$; buffer D, 50 mM

35 Tris-HCl (pH 9.0), 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 0.5 mM spermidine; buffer E, 7 M urea, 0.05% (w/v) xylene cyanol; buffer F, 5% (v/v) glycerol, 0.04% (w/v) xylene cyanol, 0.04% (w/v)

SUBSTITUTE SHEET (RULE 26)

22

bromophenol blue.

Enzymes and Antibodies. Nuclease P1 was obtained from Pharmacia Biotech. CIP was supplied by Promega. T4 Polynucleotide Kinase was purchased from New England Biolabs. AmpliTaq Polymerase was obtained from Perkin Elmer. Rabbit polyclonal anti-DNP antibody was purchased from Sigma. Anti-rabbit Ig, horseradish peroxidase-linked whole antibody was obtained from Amersham.

Analytical Methods. Melting points were determined on a Mel-Temp melting point apparatus and are uncorrected. ^{1}H NMR and ^{31}P NMR spectra were recorded on a Varian XL-300 or 500 spectrometer, respectively. Chemical shifts are reported in parts per million (δ) relative to internal tetramethylsilane or to external 85% phosphoric acid. UV spectra were measured on either a LKB Biochem Ultraspec II or a Beckman DU 7500 spectrophotometer. IR spectra were recorded with a Perkin Elmer 710B Infrared spectrometer. FAB MS were recorded on a VG analytical ZAB-E spectrometer. Unless otherwise stated, all chemical reactions were performed under a nitrogen atmosphere.

HPLC analyses were performed with a Beckman LC system which was equipped with a 126 solvent delivery module and a 168 diode array detector and was controlled with Beckman System Gold Software. A reverse phase C18 column (Beckman Ultrasphere ODS, 250 x 4.6 mm, i.d.) and guard column (Upchurch, 1 cm x 4.3 mm, i.d.) were utilized with both analytical and semi-preparative separations. Sample loops of 1 mL and 50 μL were used for semi-preparative and analytical analyses, respectively. All HPLC analyses were performed using a gradient solvent system composed of triethylammonium buffers and acetonitrile at a flow rate of 1 mL/min. Unless otherwise stated, UV absorption was monitored at a wavelength of 260 nm.

Abbreviations: DNP,2,4-dinitrophenyl; 5-IAF, 5-iodoacetamido-fluorescein; CIP, Calf intestinal alkaline phosphatase; TEAA, triethylammonium acetate; TEAB, triethylammonium bicarbonate; DMT, dimethoxytrityl; RP HPLC, reverse phase high pressure liquid chromatography; PCR, polymerase chain reaction; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; DMTCl, dimethoxytrityl

SUBSTITUTE SHEET (RULE 26)

chloride; DMT⁺BF₄, dimethoxytrityl tetrafluoroborate; BME, β–mercaptoethanol; HR FAB MS, high resolution fast atom bombardment mass spectroscopy.

Example 1. Preparation of 5-SDNP-dU-phosphoramidite

Shown in Figure 1 and described in detail below is the pathway used for the synthesis of the DNP-labeled phosphoramidite, Compound 7.

Synthesis of 3',5'-O-Diacetyl-5-thiocyanato-2'-deoxyuridine (Compound 2).

Anhydrous HOAc (150 mL) was added to a three-neck round bottom flask equipped with a vacuum adapter and teflon needle. Dry Cl_2 was slowly bubbled through the teflon needle into the reaction vessel for a brief period. A 5 mL aliquot of the Cl_2/HOAc solution was removed and subjected to iodometric titration to determine the amount of Cl_2 added (2.9 g, 41 mmol) (Flaschka et al., 1969). KSCN, previously dried overnight at 100°C, (4.4 g, 45 mmol) was added to the reaction vessel. The reaction mixture was allowed to stir at room temperature for 30 min. prior to the addition of 3',5'-0-diacetyl-2'-deoxyuridine, 3',5'-O-Diacetyl-2'-deoxyuridine was prepared by reaction of deoxyuridine with acetic anhydride in pyridine for an 82% yield: mp=106-108°C(Lin & Gao, 1983) (1.1 g, 3.5 mmol). After the mixture was stirred for an additional 1.5 hours, cyclohexene (10 mL) was added. The mixture was then stirred for 15 min. and filtered; the filtrate was concentrated under vacuum. Residual HOAc was removed azeotropically with toluene and the residue was triturated with petroleum ether. The petroleum ether was decanted and the residue subjected to column chromatographic purification [petroleum ether: EtOAc(2:1 - 1:3)] which yielded a semi-solid. Recrystallization from EtOH:EtOAc provided Compound 2 as a white solid (1.05 g; 82%):mp 154-156°C; IR(KBr) 2140cm⁻¹ (SCN stretch); ^{1}H NMR (CDCl₃) δ 9.12(bs, 1H, N₃H), 8.18 (s, 1H, H₆), 6.14 (dd, J=7.8, 5.7 Hz, 1H, H₁'), 5.13 (m, 1H, H₃'), 4.24 (m, 1H, H₅'), 4.19-4.15 (m, 2H, H₄' and H₅'), 2.32 (ddd, J= 14.4, 5.7, 2.1 Hz, 1H, H₂'), 2.13-2.07 (m, 7H, 2 CH₃ and H₂'); UV (MeOH) 270 nm.

Synthesis of 3',5'-O-Diacetyl-5-mercapto-2'-deoxyuridine, (Compound 3).

This reaction should be performed in a fume hood. EDTA (25 mL; 0.1 M, pH 7.7) and DTT (293 mg, 1.90 mmol) were added to a solution of thiocyanate 2 (207 mg, 0.561 mmol) in MeOH (30 mL). The reaction mixture was allowed to stir at room temperature for 35 min. and was then filtered. The filtrate was reduced to half by rotary evaporation, made acidic with 10% (v/v) H_2SO_4 and extracted with CH_2Cl_2. The CH_2Cl_2 extract was concentrated under reduced pressure and the residue was triturated with H_2O to provide Compound 3 as a white solid (148 mg; 77%). TLC analysis showed a yellow spot upon development with 2,2'-dithiobis(5-nitropyridine) (Grassetti & Murray, 1969): mp=159-163°C; IR(KBr) 2530 cm^{-1} (SH stretch); ^{1}H NMR (CDCl$_3$) δ 8.70 (bs, 1H, N$_3$H), 7.67 (s, 1H, H$_6$), 6.30 (dd, J=8.2, 5.8 Hz, 1H, H$_{1'}$), 5.22 (m, 1H, H$_{3'}$), 4.43-4.25 (m, 3H, H$_{4'}$ and H$_{5'}$), 3.60 (s, 1H, SH), 2.50 (ddd, J=14.4, 5.8, 2.2 Hz, 1H, H$_{2'}$), 2.21-2.10 (m, 7H, 2 CH$_3$ and H$_{2'}$); UV (0.1 M EDTA w/DTT, pH 7.76) 260, 334 nm. (The broad melting point range may indicate the presence of 5,5"-dithiobis(3',5'-diacetyldeoxyuridine). On the average, 7% disulfide analog was present in the sample as determined by UV absorbance (λmax 335 nm) in the presence and absence of DTT (Bardos & Kalman, 1966)).

Synthesis of 3',5'-O-Diacetyl-5-S-(2,4-dinitrophenyl)mercapto-2'-deoxyuridine, (Compound 4).

2,4-Dinitrofluorobenzene (57 mg, 0.31 mmol) and Et$_3$N (200 μL) were added to a solution of mercaptan 3 (86 mg, 0.25 mmol) in anhydrous CH$_3$CN (10 mL). The reaction mixture was allowed to stir at room temperature for 1.5 hours before the solvent was removed by rotary evaporation. The crude material was purified using column chromatography [Hexanes: EtOAc(4:1 to 1:1)]. The product was obtained as a yellow solid (113 mg, 89%): IR(KBr) 1530, 1340 cm^{-1} (NO$_2$ stretch); ^{1}H NMR (CDCl$_3$) δ 9.14 (d,J=2.4 Hz, 1H, H$_3$ of DNP), 8.40 (bs, 1H, N$_3$H), 8.31 (dd, J=9.1, 2.5 Hz, 1H, H$_5$, of DNP), 8.23 (s, 1H, H$_6$), 7.33 (d,J=9 Hz, 1H, H$_6$ of DNP), 6.32 (dd, J=7.8, 5.7 Hz, 1H, H1'), 5.25, (m, 1H, H$_{3'}$), 4.44-4.25 (m, 3H, H$_{4'}$ and H$_{5'}$), 2.55 (m, 1H, H$_{2'}$), 2.27 (m, 1H, H$_{2'}$), 2.13 (s, 3H, CH$_3$), 2.10 (s, 3H, CH$_3$); UV (MeOH) 264, 323 nm.

25

Synthesis of 5-S-(2,4-Dinitrophenyl)mercapto-2'-deoxyuridine, (Compound 5).

NaOMe (25% solution in MeOH; 82 μL; 0.36 mmol) was added to a solution of Compound 4 (92 mg, 0.18 mmol) in anhydrous MeOH (6 mL) and the reaction mixture was maintained at room temperature. After 2 hours, the mixture was treated with Dowex 50W x4 and filtered. The filtrate was reduced by rotary evaporation. The product was obtained as a yellow solid following column chromatographic purification of the reaction mixture [EtOAc: petroleum ether (1:4 to 1:0)] (62 mg; 81%): IR(KBr) 1530, 1340 cm^{-1} (NO$_2$ stretch); ^{1}H NMR (Cd$_3$OD) δ 9.08 (d, J = 2.4 Hz, 1H, H$_3$ of DNP), 8.85 (s, 1H, H$_6$), 8.39 (dd, J = 9, 2.4 Hz, 1H, H$_5$ of DNP), 7.64 (d, J = 9 Hz, 1H, H$_6$ of DNP), 6.33 (t, J = 6.3 Hz, 1H, H$_{1'}$), 4.45 (m, 1H, H$_{3'}$), 4.00 (m, 1H, H$_{4'}$), 3.84 (dd, J - 12, 3 Hz, 1H, H$_{5'}$), 3.75 (dd, J = 12, 3 Hz, 1H, H$_{5'}$), 2.50-2.33 (m, 2H, H$_{2'}$); UV (0.1 M TEAA, pH 7.1 with 57% CH$_3$CN) 262, 334 nm.

Synthesis of 5'-O-Dimethoxytrityl-5-S-(2,4-dinitrophenyl)mercapto-2'-deoxyuridine (Compound 6).

Anhydrous CH$_3$CN (6 mL) was added to a mixture of compound 5 (42 mg, 0.10 mmol), DMT$^+$BF$_4^-$(57 mg, 0.14 mmol) and 2,6-di-*tert*-butyl-4-methylpyridine (DBMP) (35 mg, 0.18 mmol). The resulting red solution was heated to reflux. After 1.5 hours, a second portion of DMT$^+$BF$_4^-$ (26 mg, 0.064 mmol) and DBMP (13 mg, 0.067 mmol) was added and the reaction mixture was maintained at reflux for an additional 8 hours. The CH$_3$CN was removed by rotary evaporation. CH$_2$Cl$_2$ and 5% aq. NaHCO$_3$ were added to the residue. The mixture was stirred briefly and the resulting two layers separated. The NaHCO$_3$ layer was re-extracted twice with additional CH$_2$Cl$_2$. The organic layers were combined, dried (Na$_2$SO$_4$) and reduced. The product was purified using column chromatography [CH$_2$Cl$_2$:MeOH:pyridine (99.5:0:0.5 to 98.5:2:0.5)] (27 mg: 37%); ^{1}H NMR (CDCl$_3$) δ 9.02 (d, J = 2.1 Hz, 1H, H$_3$ of DNP), 8.46 (s, 1H, H$_6$), 8.07 (dd, J = 9, 2.4 Hz, 1H, H$_5$ of DNP), 7.35-7.09 (m, 10H, H6 of DNP and Ar), 6.75-6.64 (m, 4H, Ar), 6.41 (dd, J = 9, 2.4 Hz, 1H, H$_{1'}$), 4.70 (m, 1H, H$_{3'}$), 4.16 (m, 1H, H$_4$), 3.73 (2s, 6H, 2CH$_3$), 3.38 (d, J = 7.5 Hz, 2H, H$_{5'}$), 2.64 (m, 1H, H$_{2'}$), 2.44 (m, 1H, H$_{2'}$); UV (0.1 M TEAA, pH 7.1 with 82% CH$_3$CN)

26

264, 324 nm; HR FAB MS calculated for $C_{36}H_{32}N_4O_{11}S$ (M)$^+$728.1788 found 728.1802 (1.9 ppm). (Synthesized by the method of Bleasdale et al. (Bleasdale *et al.*, 1990)).

Synthesis of 5'-O-(4,4'-dimethoxytrityl)-5-S-(2,4-dinitrophenyl)
mercapto-2'-deoxyuridine-3'-O-(2-cyanoethyl-N,N'-diisopropyl)
phosphoramidite (Compound 7).

Compound 6 (13 mg, 18 μmol) was dried under vacuum in the reaction vessel. Anhydrous CH_3CN (0.3 mL) was added and the reaction vessel was transferred to an atmosphere bag filled with dry N_2. A mixture of 2-cyanoethyl-N,N,N'N'-tetraisopropyl-phosphorodiamidite (5.5 μL, 17 μmol) and tetrazole (1.3 mg, 18 μmol) in anhydrous CH_3CN (50 μL) was added. The reaction mixture was maintained at room temperature for 30 min. An additional m i x t u r e   o f   2 - c y a n o e t h y l - N , N , N ' N ' - tetraisopropylphosphorodiamidite (2.3 μL, 7.2 μmol) and tetrazole (0.6 mg, 8.3 μmol) in anhydrous CH_3CN (25 μL) was added and the reaction mixture was stirred for 30 min. CH_2Cl_2 (20 mL) and 5% (w/v) aq. $NaHCO_3$ (15 mL) were added to the mixture; the organic layer was removed and re-extracted with water, dried (Na_2SO_4) and evaporated. The residue was subjected to column chromatographic purification [CH_2Cl_2:MeOH:pyridine(99.5:0:0.5 to 98.5:1.0:0.5)] to obtain Compound 7 (8.6 mg, 52%). The yellow product was analyzed by analytical RP HPLC using a gradient of buffer A and acetonitrile [70% (v/v) CH_3CN to 80% in 3 min.; 80% CH_3CN to 100% in 15 min.]. UV absorption was monitored at wavelengths of 254 and 330 nm. The diastereomers eluted at 11.09 and 12.49 min. UV scans of the diastereomers showed absorption peaks at 264 and 327 nm. Phosphorus NMR indicated the presence of two diastereomers [^{31}P NMR ($CDCl_3$) δ 149.12, 149.27 ppm].

Example 2.   Automated Synthesis of an Oligonucleotide
Containing 5-SDNP-dU

Synthesis and Purification of Oligonucleotides.

Oligonucleotides were synthesized on an Applied Biosystems 392 DNA Synthesizer using the standard β-cyanoethyl-protected phosphoramidite method. A series of three modified oligonucleo-

SUBSTITUTE SHEET (RULE 26)

tides were prepared on a 40 nmole scale using a Trityl On method. The thiol-protected analog, Compound 7, was manually dissolved into anhydrous acetonitrile at a concentration of 0.027 M and the solution was attached to the spare port on the DNA synthesizer. As the concentration of Compound 7 was lower than the concentration of the commercial phosphoramidites used in these syntheses (0.05 M), the coupling time for the modified analog was increased by 120 seconds. Syntheses of the corresponding unmodified oligonucleo-tides were also performed. The sequences of the oligonucleotides are shown in Table 1.

The oligonucleotides were removed from the column by treatment with concentrated NH₄OH(1 mL). Deprotection of the exocyclic amines was then accomplished by treatment of the oligonucleotides with an additional 1 mL of fresh concentrated NH₄OH at room temperature for 44 hours. The NH₄OH solution was concentrated using a Savant Speed Vac without heat. Small portions of Et₃N were periodically added to the concentrating solution to maintain a basic pH. The crude oligonucleotides were re-dissolved in 10 mM TEAB (pH 7.1) and purified as the DMT ethers using semi-preparative RP HPLC. Two different gradients of buffer B and acetonitrile were utilized. The shorter oligonucleotides (7 and 22 nucleotides) were purified with an elution gradient of CH₃CN from 5% (v/v) to 30% in 15 min. then 30% to 100% in 5 min. The 24-mer oligonucleotides were purified with an elution gradient of CH₃CN from 5% to 30% in 18 min. then 30% to 100% in 5 min. The appropriate fractions were evaporated to dryness using a Jouan Concentrator without heat. Following purification of the oligonucleotides, the DMT group was removed by treatment with 80% (v/v) HOAc (200 µL) at room temperature for 20 min. The oligonucleotides were ethanol precipitated and re-dissolved into 300 µL of nanopure water and stored at -20°C. The purity of the oligonucleotides was analyzed by analytical RP HPLC (Figure 2A, Oligonucleotides A and B shown). Elution of the oligonucleotides was performed using a linear gradient of CH₃CN (5 to 20% in 15 min.) with buffer B. The purified oligonucleotides were also analyzed by 25% polyacrylamide-7 M urea gel (acrylamide/methylene bisacrylamide = 19/1) (Figure 3A, Oligonucleotides C-F shown).

28

Enzymatic Digestion with Nuclease P1 and CIP. The oligonucleotides A and B were subjected to enzymatic digestion. The pure oligonucleotides were incubated with Nuclease P1 (2 U) at 37 °C for 6 hours in a total volume of 50 μL of buffer C.

5   Five microliters of 10X buffer D and 43 μL of H₂O were added to the reaction. CIP (2 U) was added and the pH of the reaction was adjusted with 5% (v/v) NaOH (0.8 μL). The reaction mixture was incubated overnight at 37°C. To recover the nucleosides, 11 μL of 3 M NaOAC, pH 5.3 and 250 μL of absolute EtOH were added

10   to the reaction mixture. The mixture was stored at -80°C for 4 hours and then centrifuged at 14000 rpm for 20 min. The supernatant fraction was removed and concentrated. The resulting residue was re-dissolved into 35 μL of buffer A and analyzed by analytical RP HPLC (Figure 2B). The elution

15   gradient consisted of buffer A with an increasing concentration of CH₃CN (0% to 5% in 5 min., 5% to 6% in 15 min., 6% to 50% in 5 min. and finally 50% to 80% in 15 min.). The UV absorption was monitored at 260 and 330 nm. Commercial nucleosides and monophosphates of dA, dC, dG and T were utilized as standards.

20   Synthetic 5-S-(2,4-dinitrophenyl) mercaptodeoxyuridine (Compound 5) was also utilized as a standard. An enzymatic digestion without oligonucleotide was performed to identify background peaks arising from the buffer and enzymes.

Deprotection and alkylation of a DNP-modified oligonucleotide.

25   Oligonucleotide F (100 pmol) was radiolabeled by reaction with T4 Polynucleotide Kinase using [γ-³²P]ATP. The labeled oligonucleo-tide F was isolated by ethanol precipitation and re-dissolved into 30 μL H₂O. The oligonucleotide (4.5 μL) was treated with 1.4 M BME either at 45 °C for 4 hours or at room

30   temperature for 12 hours in a 20 μL solution of 20 mM Tris-HCl, pH 8.2. Following this deprotection to remove the DNP group, half of the reaction mixture was precipitated with ethanol while the other half of the mixture was subjected to alkylation conditions with 5-IAF. The corresponding unmodified

35   oligonucleotide E was treated identically to serve as a control.
        To recover the deprotected oligonucleotides by ethanol precipitation, 0.36M NaOAc, pH 5.6 (100 μL), 1M MgCl₂(1.2 μL) and

SUBSTITUTE SHEET (RULE 26)

nanopure water (10 μL) were added to the reaction mixture (10 μL) to provide a final volume of 121 μL. Absolute ethanol (300 μL) was added and the mixture was cooled at -80°C. The oligonucleotides were pelleted by centrifuging at 14000 rpm for 20 min. The supernatant was removed and the pellet was washed with cold 95% (v/v) ethanol. The mixture was chilled for 20 min. in a dry ice/ethanol bath and spun at 14000 rpm for 20 min. Following the removal of the supernatant fraction, the recovered oligonucleotides were dissolved into 20 μL of nanopure water. Mixtures of oligo-nucleotides (4.5 μL), 100 mM DTT (0.5 μL) and buffer E (5 μL) were prepared and analyzed by electrophoresis on a 25% polyacrylamide-7 M urea gel (acrylamide/methylene bisacrylamide = 19/1) (Figure 3B).

Alkylation of the deprotected thiol moiety was accomplished by the addition of 15 μL of 20 mM Tris-HCl (pH 4.8) to 10 μL of the deprotection reaction (pH 8.2). Following this addition, the pH of the alkylation reaction was 7 (as judged by pH paper). One microliter of 10mM 5-IAF in DMF was added and the reaction mixture was maintained at room temperature overnight. The oligonucleotides were recovered by ethanol precipitation and re-dissolved into 20 μL nanopure water. The alkylated oligonucleotides (4.5 μL) were mixed with 100 mM DTT (0.5 μL) and buffer E (5 μL) and analyzed by electrophoresis on a 25% polyacrylamide-7M urea gel (acrylamide/methylene bisacrylamide = 19/1) (Figure 3B).

DNA Synthesis via PCR using a DNP-modified oligonucleotide.
PCR reactions were performed on a Perkin Elmer Cetus Gene Amp PCR System 9600. The DNA fragment containing the bacteriophage lambda PR' promoter, 6S gene, and tR' terminator, with the lambda qut site (Yang et al., 1987) was amplified from plasmid pHA100 (Zhang & Hanna, 1995) using either oligonucleotide E or F and the universal T7 primer (Stratagene). The DNA product was purified by excision of the band from an agarose gel and extraction of the DNA using the commercially available GENECLEAN kit (BIO 101 Inc., La Jolla, CA). The PCR products were dissolved into equal volumes of buffer F and analyzed by electrophoresis on a 1.5% agarose gel. Detection of

30

DNA was achieved by addition of ethidium bromide to the electrophoresis buffer. (Figure 4A).

Immunodetection of DNP-labeled DNA.

Dot blot experiments were performed manually. The DNP-labeled
5   PCR product and the corresponding unlabeled DNA fragment (1 pmole each) were in 10X SSC solution (1.5 M NaCl, 0.15 M Na₃Citrate, pH7). Samples (13 µL) were applied to a pre-wetted charged nylon membrane (Zeta Probe GT; Biorad). The membranes were air dried, rinsed briefly with 2X SSC and heated at 80°C
10   under vacuum for 30 min. The membranes were blocked by overnight incubation at 4°C in TBS-T (20 mM Tris HCl, 500 mM NaCl, 0.4% (v/v) Tween 20, pH 7.6) containing 10% (w/v) dried milk. The membranes were then washed with TBS-T two times briefly, one time for 15 min., and two times for 5 min. The
15   membranes were incubated with a 1:1500 dilution of anti-DNP antibodies for 1 hour, washed as described above and then incubated with a 1:10,000 dilution of horseradish peroxidase-linked whole anti-rabbit Ig for 1 hour. Membranes were then washed one time for 15 min. and 4 times for 5 min. with TBS-T.
20   Detection of the antibodies was carried out using the ELC chemiluminescence method (Amersham). The membranes were incubated with the detection agents for 1 min. before exposure to x-ray film for 2 min. (Figure 4B).

RESULTS

25   The synthesis and characterization of a novel deoxyuridine phosphoramidite (Compound 7) containing a protected thiol group at the 5 position of the mercaptopyrimidine ring is described (Figure 1). This DNP-labeled analog was site-specifically incorporated internally into a series of oligonucleotides.
30   Significant differences in overall synthetic yields were not observed between oligonucleotides made with normal deoxynucleoside phosphoramidites and those which contained the analog, indicating that the modified phosphoramidite is a suitable reagent for automated oligonucleotide synthesis.
35   Enzymatic digestion of the DNP-labeled oligonucleotide established the stability of the modified pyrimidine to the

conditions for chemical synthesis and purification of oligonucleotides (Figure 2B). Selective deprotection of the 5-thiol moiety with BME and subsequent modification of the unmasked thiol with 5-IAF were verified by gel electrophoresis
5   of the oligonucleotides (Figure 3B). Incorporation of the DNP-labeled oligonucleotide into double-stranded DNA was achieved using PCR. Isolation and characterization of the PCR products included gel electrophoresis and immunodetection with anti-DNP antibodies (Figure 4A and B).

10   **Synthesis.**
         Synthesis of the deoxyuridine phosphoramidite analog (Compound 7) was carried out following the sequence of reactions outlined in Figure 1. Compounds 2-7 have not been previously described in the literature. The method chosen for
15   incorporation of the thiol group required thiocyanation of the pyrimidine ring. Initially, this reaction was performed on deoxyuridine and the reaction conditions utilized were similar to those previously described (Bradley & Hanna, 1992; Nagamachi et al., 1974; Torrence et al., 1968). Although literature
20   reports indicate yields of 54% for the thiocyanation of deoxyuridine, this was difficult to achieve. The yield for this reaction varied considerably and was at times as low as 9%. A major side product in this reaction was isolated and identified by NMR and IR as 5'-O-acetyl-5-thiocyanatodeoxyuridine
25   indicating that along with thiocyanation, acetylation of the sugar ring had occurred.
         To circumvent the problems associated with the reaction described above, 3',5'-O-diacetyl-2'-deoxyuridine (Compound 1) was utilized as the starting material. Protection of the sugar
30   hydroxyls by acetylation prior to thiocyanation results in higher yields (95%) for the reaction (Nagamachi et al., 1974). However, removal of the acetyl protecting groups could not be accomplished using standard basic conditions required for deesterification without destruction of the thiocyanato group.
35   Therefore, a synthetic approach was developed which required reduction of the thiocyanato compound to the mercapto analog prior to removal of the acetyl groups (Figure 1). Thiocyanation

of the 3',5'-O-diacetyl-2'-deoxyuridine (1) was achieved following the literature procedure described above except that the crude material was extracted with petroleum ether to remove nonpolar impurities prior to column chromatographic purification. The pure product was obtained in 82% yield as a white solid.

Reduction of the thiocyanato derivative 2 to the mercapto analog (Compound 3) was achieved by reaction with DTT in a solvent mixture of MeOH and 0.1 M EDTA (pH=7.8). This reaction is a variation of a reported procedure (Lin et al., 1988) for the reduction of 3'-azido-2',3'-dideoxy-5-thiocyanatouridine; this mercapto analog however was not isolated. Purification of Compound 3 was achieved by acidification of the reaction mixture followed by concentration of the mixture. The residue was extracted with CH_2Cl_2 and H_2O. The presence of the acetyl groups insured the product was organic solvent soluble. In order to remove residual DTT, the organic layer was reduced and the residue triturated with water to crystallize the product as well as to dissolve residual DTT. Compound 3 was obtained as a white solid in 75% yield.

The mechanism for the reduction of the thiocyanato requires displacement of the cyano group by DTT. Since no attempt was made to remove the cyanide ion before acidification of this reaction mixture, every step of the isolation of Compound 3 from acidification to trituration with H_2O was performed in a fume hood. This synthesis of Compound 3 was designed for the preparation of both 5-alkyldithio and 5-alkylthio or 5-arylthiopyrimidine analogs. To prevent degradation of the 5-alkyldithio compounds by DTT, purification of Compound 3 was necessary. This is not the case with formation of 5-alkylthio pyrimidines (Lin et al., 1988) and this purification step has been eliminated for Compound 4 resulting in a slight decrease in yield.

The DNP protecting group has been employed in oligonucleotide synthesis as a protecting group for 6-mercaptopurine and 6-thioguanine phosphoramidites (Xu et al., 1992a; Xu et al., 1992b) and as a functional moiety on a non-nucleoside phosphoramidite (Grzybowski et al., 1993). These

studies indicated the DNP group was stable during automated DNA chemical synthesis. Conversion of mercaptan 3 into the 2,4-DNP analog 4 was achieved by reaction with 2,4-dinitrofluorobenzene in anhydrous CH₃CN with Et₃N present as a catalyst (Xu et al.,

5   1992b). The product was isolated as a yellow solid in 88% yield. Attachment of the thiol protecting group prior to removal of the acetyl moieties insured that the product was soluble in organic solvents, which aided in the purification of the material. More importantly, protection of the sulfur

10   inhibits the oxidation of the analog to the disulfide compound, a potential reaction of 5-mercaptopyrimidines (Kalman & Bardos, 1967).

Another method of preparing 5-arylthiol ethers of deoxyuridine has been previously described. Bergstrom et al.

15   has prepared a series of alkyl and aryl mercaptodeoxyuridine analogs using a palladium-mediated reaction between 5-(chloromercuri)-2'-deoxyuridine and the appropriate disulfide (Bergstrom et al., 1991). Although this method offers a simple route to arylthiol ether analogs it was not selected for study.

20   Development of a more versatile intermediate such as Compound 3 which can be easily converted into various types of analogs including disulfides was desired for future studies. Additionally, the synthetic utility of the method has not been fully evaluated and therefore formation of the desired

25   dinitrophenyl analog (Compound 5) was uncertain.

Deesterification of nucleosides is commonly performed under basic conditions using NH₃ or NaOMe (Kini et al., 1989; Trivedi & Bruns, 1989). Literature reports suggest the DNP group is marginally stable to nucleophiles such as NaOMe; the stability

30   of this group being strongly dependent on the reaction conditions (Greene, 1981). Successful removal of the acetyl groups was achieved by reaction of the diacetyl analog (Compound 4) with NaOMe in MeOH without an apparent loss or modification of the DNP group as judged by NMR and IR. Following the

35   deacetylation, the diol (Compound 5) was isolated as a yellow solid in 81% yield. An initial attempt to deacetylate Compound 4 under mild reductive methods was performed (Brown et al., 1982; Soai & Ookawa, 1986). However, reaction with LiBH₄ in a

SUBSTITUTE SHEET (RULE 26)

mixture of Et_2O and MeOH did not result in the formation of Compound 5, possibly due to the poor solubility of the starting material.

Incorporation of the DMT protecting group at the 5' position of nucleosides is typically accomplished by reaction with DMTCl in the presence of pyridine or dimethylpyridine (Chaudhary & Hernandez, 1979). However, as utilization of these conditions with other 5-thiol modified nucleosides was not successful (unpublished results), this reaction was not attempted for the synthesis of Compound 6. Conversion to the 5'-dimethoxytrityl analog (Compound 6) was achieved by reaction of Compound 5 with DMT^+BF_4- in the presence of DBMP in CH_3CN at reflux. This reagent, DMT^+BF_4, has only recently been utilized in the preparation of 5' protected nucleosides (Bleasdale et al., 1990; Lakshman et al., 1992).

The proposed mechanism for dimethoxytritylation with DMTCl and pyridine requires formation of a dimethoxytrityl pyridinium salt which then undergoes reaction with the 5' hydroxyl group (Chaudhary & Hernandez, 1979). Formation of the pyridinium salt is considered the limiting step. The reagent DMT^+BF_4 serves as a source for the dimethoxytrityl cation. By preforming the cation prior to reaction, the slow step in the dimethoxytritylation has been eliminated and the ether should be more easily formed. Although Compound 6 was obtained using this reagent, the reaction was performed under reflux conditions. It is unclear why the reaction requires such stringent conditions but presumably steric hindrance plays some part. Modification of other bases has been reported to inhibit dimethoxytritylation possibly due to steric hindrance (Cosstick & Douglas, 1991).

Silica gel column chromatography was employed to isolate Compound 6. Due to the acidic nature of the silica gel, a small percentage of pyridine was added to the eluting solvent to prevent loss of the acid sensitive DMT group. Initially, Et_3N was used in the solvent system; however triethylammonium salts of Compound 6 were isolated as determined by NMR. This was not observed during the synthesis of the standard 5'-dimethoxytrityl-2'-deoxyuridine. Switching to pyridine, a weaker base, in the eluting solvent system eliminated this

35

problem.   A satisfactory high resolution FAB mass spectra was obtained for Compound 6.

The Compound 6 was isolated as a yellow solid in yields ranging from 35 to 48%, making formation of the dimethoxytrityl ether the lowest yielding reaction in this synthetic sequence. However, approximately 20% of the starting material, Compound 5, was converted into the 3',5'-bis(dimethoxytrityl) analog, which was isolated and converted back into 4 for recycling into the synthesis.

The final reaction in the preparation of Compound 7 is the incorporation of the phosphoramidite moiety at the 3' position of the nucleoside. This is commonly accomplished by reaction of the appropriately protected nucleoside with 2-cyanoethyl-N,N-diisopropyl chlorophosphoramidite in the presence of the N,N-diisopropylethylamine (DIEPA) (Sinha et al., 1984). Utilization of these conditions did not result in the desired product and starting material was re-isolated. However, a standard reaction with 5'-dimethoxytrityldeoxyuridine was successful. A second method for incorporation of the phosphoramidite moiety which utilizes the reagent 2-cyanoethyl-N,N,N,'N'-tetraisopropylphosphorodiamidite in the presence of tetrazole in CH$_3$CN was evaluated (Barone et al., 1984). Unlike the basic reaction with DIEPA, these slightly acidic conditions provided the DNP-labeled phosphoramidite, Compound 7. RP HPLC analysis of Compound 7 indicated the presence of a pair of closely eluting peaks (11.0 and 12.4 min) representative of the expected 1:1 mixture of diastereomers. The overall purity of the sample was determined to be greater than 96% and no starting material was observed in the chromatogram. HPLC data was quantified using a simple area percent method. Similar results were observed using silica gel TLC analysis (CH$_2$Cl$_2$:MeOH:Et$_3$N;96:3:1). The phosphorous NMR spectrum showed two characteristic peaks associated with the diastereomeric phosphoramidites (149.12 and 149.27 ppm). An additional peak appeared in the NMR at 14.3 ppm. Based on the chemical shift of this peak, the impurity is presumed to be a hydrolysis product (Scremin et al., 1994). No attempt was made to remove this material prior to oligonucleotide synthesis as it is not reactive in the

36

oligonucleotide chemistry. A high resolution FAB mass spectrum could not be obtained for Compound 7 due to its instability during analysis.   This was not surprising as the standard deoxyuridine phosphoramidite displayed a high degree of

5   instability during high resolution FAB MS analysis; however, a small molecular ion peak was identified in the spectrum. Evaluation of Compound 7 using low resolution mass spectroscopy showed a fragment (m/z=407) consistent with the loss of the DMT and the phosphoramidite moieties.   As a similar fragmentation

10   pattern was observed for the standard deoxyuridine phosphoramidite under these experimental conditions, the information supports the proposed structure.

Syntheses and purification of oligonucleotides.

Internal incorporation of the phosphoramidite analog,

15   Compound 7, into series of oligonucleotides has been accomplished using an automated DNA synthesizer. The natural oligonucleotides with corresponding sequences were prepared as standards.   The sequences for the oligonucleotides which are listed in Table 1 were chosen for two reasons: the 7-mer was

20   designed for characterization purposes while the 22-mer and the 24-mer were developed to study the DNA-protein interactions during transcription of the bacteriophage lambda 6S gene. The concentration of the phosphoramidite solution of Compound 7 was diluted twofold as compared to the normal phosphoramidites (dA,

25   dG, T and dC) utilized in these DNA syntheses. This dilution was performed to conserve material. The coupling time for the reaction involving the modified phosphoramidite was increased by 120 seconds.   Under these reaction conditions, the coupling yield for the modified phosphoramidite was >96% which was

30   comparable to the coupling yields of the other standard phosphoramidites.   The average stepwise yields for all six oligonucleotide syntheses were determined automatically by the ABI synthesizer using a trityl cation assay. These yields are listed in Table 1.

35   Selective removal of the exocyclic amine protecting groups in the presence of the DNP moiety was achieved by treatment with NH$_4$OH at room temperature for 44 hours.   Typically, this

37

deprotection is performed at elevated temperatures of 55°C for
8 hours (ABI technical bulletin #13, 1987). However, studies
involving reaction of the DNP-labeled nucleoside 5 with conc
NH₄OH at 55°C indicated the analog was not stable to these
5   conditions but was stable at room temperature (data not shown).
As successful removal of standard exocyclic amine protecting
groups at room temperature has been reported for other
oligonucleotides containing modified phosphoramidites, these
conditions were utilized (Sheardy & Seeman, 1986). Although
10   these initial studies were performed with the standard exocyclic
amine protecting groups, phosphoramidites containing a wide
variety of protecting groups which can be efficiently removed
under mild conditions are now commercially available. These
will be investigated for compatibility with phosphoramidite
15   Compound 7.

The crude oligonucleotides were purified as the DMT ethers
by RP HPLC. Removal of the DMT group under standard acidic
conditions did not affect the DNP moiety. Following removal of
the DMT group, the oligonucleotides were analyzed for purity by
20   both analytical RP HPLC and gel electrophoresis as shown in
Figures 2A and 3A respectively. The oligonucleotides showed one
major product in both analyses. The retention times of the DNP-
labeled oligonucleotides were greater than the corresponding
natural oligonucleotides in the HPLC analyses due to the
25   lipophilic nature of the dinitrophenyl group. The RP HPLC
retention times for the purified oligonucleotides are listed in
Table 1. The overall yields of the pure material which were
calculated by dividing the optical density of pure
oligonucleotide by the optical density of crude DMT-labeled
30   oligonucleotides times 100% are found in Table 1 (Kuimelis &
Nambiar, 1994). The yields for the corresponding normal and
DNP-labeled oligonucleotides were comparable.

Enzymatic digestion of the heptanucleotide.

To establish that the DNP group is stable to the conditions
35   of automated DNA synthesis, deprotection, and purification, a
DNP-labeled oligonucleotide was enzymatically digested into its
corresponding nucleosides and evaluated by RP HPLC. In order to

SUBSTITUTE SHEET (RULE 26)

38

observe the single DNP-labeled nucleoside present in the enzymatic digest, the short chain oligonucleotide B was used for this analysis. The sequence chosen for this heptanucleotide did not contain deoxycytidine. Deoxycytidine co-elutes with 5-
5   mercapto-deoxyuridine, a potential side product arising from the loss of the DNP group, in the gradient system utilized in the analysis. Deoxycytidine was therefore omitted from the sequence to eliminate uncertainty in identifying the nucleoside peaks. 5-Mercaptodeoxyuridine was synthetically prepared in this lab by
10   reduction of 5-thiocyanatodeoxyuridine (Nagamachi et al., 1974).

Initial degradation of the oligonucleotide B and the corresponding normal oligonucleotide A was attempted by concurrent digestion with snake venom phosphodiesterase (SVP) and CIP (Kuimelis & Nambiar, 1994). This method proved useful
15   in the digestion of the natural oligonucleotide but did not appear to fully degrade the DNP-labeled material. This is consistent with other studies in which SVP has failed to function with some modified oligonucleotides and nucleotides (Gao et al., 1992; Hanna et al, 1989; Hanna et al., 1993).
20   Substitution of the exonuclease SVP with an endonuclease Nuclease P1 resulted in complete digestion. Due to a difference in the pH requirements of these enzymes for activity, this digestion was accomplished in two consecutive enzymatic reactions. First, Nuclease P1 cleaved the phosphodiester bonds
25   to provide the 5'-monophosphates which were then subjected to treatment with CIP. The resulting nucleosides were isolated and analyzed by RP HPLC. The results are shown in Figure 2B.

This enzymatic degradation provided the correct ratios of unmodified and modified nucleotides. For oligonucleotide A, the
30   ratio of dA, dG and T was 1.9:2.1:2.9. For oligonucleotide B, the nucleoside 5 is also expected in the reaction mixture providing a ratio of dA, dG, T and 5 of 1.9:2.1:1.9:1.1. Chemically synthesized Compound 5 was utilized as a standard. Both the enzymatically obtained and the chemically obtained
35   nucleoside Compound 5 eluted with the same retention time during RP HPLC analysis. In addition, these compounds provided identical UV spectra with absorption peaks appearing at 262 and 334 nm. Finally, no unidentifiable peaks were observed in the

HPLC chromatograms of the enzymatic digests and only one peak representing nucleoside Compound 5 appeared in the HPLC chromatogram monitored at 330 nm. In this analysis, absorption at 330 nm is associated with the presence of either the dinitro-phenylthiol ether or the 5-mercaptopyrimidine moiety. If the DNP-labeled analog had undergone modification during automated DNA synthesis, deprotection, or purification, extraneous peaks were predicted to appear in this chromatogram.

Deprotection and alkylation of a DNP-modified oligonucleotide.
Removal of the DNP moiety was accomplished by reaction of oligonucleotide F with 1.4 M BME in Tris buffer pH 8 (Shaltiel, 1967). The reaction was performed under two different reaction conditions: room temperature for 12 hours and 45°C for 4 hours. Oligonucleotide E was reacted under the same conditions to serve as a standard. No attempt was made to isolate the deprotected oligonucleotide prior to modification with 5-IAF. The pKa of BME is 9.6 while the pKa of 5-mercaptodeoxyuridine is 5.0 (Bardos & Kalman, 1966). Modification of a thiol by reaction with haloacetamido compounds such as 5-IAF requires formation of the thiolate anion. By lowering the pH of the reaction mixture, the selectivity for modification of the oligonucleotide is enhanced over that of BME, a potential side reaction. The oligonucleotide products in both the deprotection and modification reactions were isolated by ethanol precipitation and analyzed by electrophoresis using a 25% polyacrylamide-7 M urea gel. The results are shown in Figure 3B. Lanes 1 through 4 represent reactions utilizing oligonucleotide E. No difference is observed in these four lanes indicating the natural oligonucleotide structure is stable to the conditions needed for deprotection and alkylation of the thiol. Lanes 5 and 7 represent reactions of oligonucleotide F with BME at room temperature and 45°C respectively, while lanes 6 and 8 depict the corresponding alkylation reactions which were carried out at room temperature. No apparent mobility shift is observed between the natural oligonucleotide (lanes 1-4) and lanes 5 and 7, which is consistent with our observations that a 20-mer RNA containing 5-SH-UTP co-migrates in this gel system with 20-mer

40

containing no analog (He et al., 1995). In contrast, a
substantial difference is observed following modification with
fluorescein, consistent with other reports of gel mobility
shifts of fluorescein-labeled oligonucleotides (Clegg et al.,
5   1992; Mergny et al., 1994). In fact, the selection of 5-IAF for
the alkylation experiments instead of a thiol-modifying
photocrosslinking reagent such as p-azidophenacyl bromide (APB)
was due in part to the ease in which the reaction results could
be monitored by gel electrophoresis. The conditions for
10   alkylation of the oligonucleotide with 5-IAF however should
prove useful with other haloacetyl derivatives such as APB. We
have specifically labeled the thiol group in an RNA molecule
containing 5-SH-UMP (He et al., 1995) with APB. We have also
attached APB to a DNA molecule generated by PCR with a
15   oligonucleotide containing (compound 7), after BME removal of
the DNP group (not shown).

As the mobility shift associated with the DNP label is
small (Figure 3A), it is possible that a mixture of DNP-labeled
and 5-mercapto-labeled oligonucleotides may appear as one band
20   on a 25% polyacrylamide-7 M urea gel. Therefore, one can not
determine the extend to which deprotection has occurred.
However, the appearance of a new band in lanes 6 and 8 (Figure
3B) with the subsequent disappearance of the band seen in lanes
5 and 7 indicates that the oligonucleotide underwent
25   deprotection and alkylation. Although deprotection does occur
with time at room temperature, the reduced time required to
deprotect at 45°C makes this the preferred method.

PCR reaction using a DNP-modified oligonucleotide.

The DNP-labeled oligonucleotide was incorporated into
30   double stranded DNA using PCR. Plasmid pHA100 served as the
template for this reaction. A 492 base pair segment of this
plasmid was amplified using the universal T7 primer and either
oligonucleotide E or F. The products isolated from the reaction
were analyzed on a 10% polyacrylamide-7 M urea gel (data not
35   shown) and a 1.5% agarose gel (Figure 4A). Differences in
mobility were not observed between the standard PCR product and
the DNP-labeled PCR product in these gel analyses and no

SUBSTITUTE SHEET (RULE 26)

difference in yield was observed.

Immunodetection of DNP-labeled oligonucleotides and DNA.

To insure the stability of the DNP protecting group during
the PCR reaction, the products described above were subjected to
5   immunodetection studies using rabbit polyclonal anti-DNP
antibodies.   The natural PCR product served as a negative
control.   The DNP group was detected using a chemiluminescence
method and only the DNP-labeled nucleic acid provided a signal
(Figure 4B).

10   Example 3:   Analysis of Protein-DNA Interactions
                        Utilizing 5-S-DNP-dU

Photochemical crosslinking is a powerful technique for
characterization of both RNA-protein and DNA-protein
interactions in nucleoprotein complexes.   Photochemical
15   crosslinking "traps" weak or transient associations which might
not survive isolation procedures such as immunoprecipitation,
gel filtration, or filter binding.   There are a number of
approaches which can be taken in photochemical crosslinking.
One involves direct irradiation of a nucleoprotein complex with
20   short wavelength ultraviolet light. This method relies upon the
direct excitation of nucleotides or amino acids to generate
chemically reactive species.   Such nonspecific labeling can
reveal whether a protein-nucleic acid interaction exists, but
there are several problems associated with this approach. These
25   include low crosslinking yields, the inability to incorporate a
crosslinking group at a specific site in the nucleic acid, and
degradation of some proteins during irradiation.   A second
approach involves the use of nucleotide analogs modified with
photoreactive crosslinking groups (reviewed in M. Hanna, 1989,
30   Methods in Enzymology).   These groups are chemically inert in
the absence of light but are converted to chemically reactive
species upon irradiation. Often this can be achieved with long
wavelength ultraviolet or visible light, resulting in less
nicking of protein or nucleic acid.   Molecular contacts can be
35   identified at the level of specific nucleotides in nucleic acids
and amino acids in proteins.   One can therefore gain structural

SUBSTITUTE SHEET (RULE 26)

42

information about macromolecular complexes which cannot be obtained by most other biochemical approaches.

For such mechanistic studies a variety of nucleotide analogs have been developed and utilized (M. Hanna 1989, Methods in Enzymology; Bradley and Hanna, 1992). These analogs are either tagged with a photoreactive crosslinking group or contain functional groups that can be tagged with crosslinkers after incorporation into DNA or RNA. These analogs can be incorporated into nucleic acids, either enzymatically or chemically, to analyze molecular interactions in protein-nucleic acid complexes. Upon photoactivation, analog-tagged nuecleic acids become covalently attached to adjacent macromolecules (protein, DNA, RNA) with which they have direct interactions. Therefore, nucleic acid binding domains in complexes can be characterized, nucleic acid binding proteins in extracts can be identified, and determinants involved in these specific interactions can be characterized at the level of individual nucleotides and amino acids. This methodology allows mechanistic questions about the way that specific protein-DNA or protein-RNA interactions regulate gene expression to be asked. Herein described are the approaches used to characterize the protein-DNA interactions in the E. coli transcription complex which forms on the bacteriophage lambda $P_{R'}$ promoter.

Preparation of Site-Specifically Modified DNA Oligonucleotides

The method described here is used to probe for interactions between a specific nucleotide in DNA and a DNA-binding protein. In this case, interactions between a specific nucleotide on the non-template strand of the $P_{R'}$ promoter are sought. The chemical approach to preparation of site-specifically modified DNA involves incorporation of a modified nucleotide during the automated synthesis of an oligonucleotide. Because the conditions used in automated oligonucleotide synthesis are rather harsh, the direct incorporation of a photocrosslinking azide group by this method is not possible, and a means for the post-synthetic incorporation of the azide group is required. 5-S-DNP-dU was incorporated into a DNA oligonucleotide, as described herein. For analysis of protein interactions with the

BNSDOCID: <WO____9803532A1_IA>

-12 position in the nontemplate strand of the $p_{R'}$ promoter, an oligonucleotide with the sequence 5'-5D-DNP-dU-AA ATT TGA CTC AAC GAT GGG (SEQ. ID. NO:1) was synthesized on an Applied Biosystems 392 DNA Synthesizer using the standard β-cyanoethyl-protected phosphoramidite method by attaching the analog to the fifth substrate port. After synthesis, removal of the oligonucleotide from the column, and removal of the exocyclic amine protecting groups, the oligonucleotide was ethanol precipitated, dissolved in water and stored at -20°C.

Analysis of 5-S-DNP-dU Effects on Hybridization Properties of ODNs

To evaluate the effect of incorporation of one or more 5-SDNP-dU analogs into oligodeoxynucleotides (ODNs) on their hybridization properties, the ODNs shown in Table 2 were synthesized. The observed melting temperatures (T_m) showed that incorporation of one DNP analog decreased the melting temperature by only 1.8°C (hybrid V), compared to the normal DNA hybrid (I), and substitution with two DNP analogs caused a decrease of only 4.4°C (hybrid VI). Therefore, the effect of a single substitution is less than that caused by a single mismatch involving unmodified nucleotides (hybrids II, III and IV vs I), and the effect of two modified analogs is still less than that of a single T-C mismatch (hybrid IV). Mismatches of 2 or more are commonly used for site-directed mutagenesis, which involves hybridization of the oligonucleotides to single-stranded DNA. Substitution with the DNP analog should therefore work as well, if not better. The use of ODNs containing one or more 5-SDNP-dU analog in assays requiring specific hybridization to a complementary strand is therefore quite feasible.

By comparison, oligodeoxynucleotides (30-mers) containing a previously described nucleobase fluorescent derivative, 5-Amido-(Carboxyfluorescein)-2'-deoxyuridine, caused decreases in melting temperature of 2°C for ODNs with one nucleobase substitution, 6°C for ODNs with two nucleobase substitutions, and 12°C for ODNs containing three nucleobase substitutions (Jadhav et al. 1997). A more serious effect on the hybridization properties was observed for another previously

44

described analog, 2-thiodeoxyuracil. Placement of even a single
2-thiodeoxyuracil analog into oligodeoxynucleotides (14-mers)
caused a decrease in T_m of 3-4°C (Kuimelis and Nambiar, 1994).

Table 2. T_m values for native and modified ODNs

ODN 1 (SEQ. ID. NO:2)	5' GAC	VCA	ACG	AWG	GGX	TAA	TYC	G	3'
ODN 2 (SEQ. ID. NO:3)	3' CTG	AGT	TAC	TZC	CCA	ATT	AAG	C	5'

ODN	1	1	1	1	2		
Hybrid	V	W	X	Y	Z	T_m(°C)	Drop in T_m[a]
I	T	T	T	T	A	67.3	0
II	T	T	T	T	G	64.5	2.6
III	T	T	T	T	T	63.6	3.5
IV	T	T	T	T	C	61.5	5.6
V	T	5[b]	T	T	A	65.5	1.8
VI	T	5	5	T	A	62.9	4.4
VII	T	5	T	5	A	62.5	4.8
VIII	5	5	T	5	A	58.0	9.3

a - as compared to unmodified duplex (Hybrid I)

b - 5 refers to 5-S-DNP-dU

Attachment of a Crosslinking Group at a Specific Position in DNA

The analog-modified oligonucleotide (100 pmol) was radiolabeled with T4 polynucleotide kinase using [γ^{32}P] ATP (Figure 5, Step 1), isolated by ethanol precipitation and re-dissolved in water. The presence of the 5-S-DNP group on the 5' terminal nucleotide did not interfere with recognition and modification of the 5' OH group by kinase. The oligonucleotide was then incorporated into double-stranded DNA by either PCR (not shown) or by primer-extension. For primer-extension using a single-stranded phagemid DNA template, the oligonucleotide (20 pmoles) was annealed to the template (1 pmol) by heating the mixture to 75°C, cooling to 65°C and then placing the mixture on ice (Step 2). The oligonucleotide was then extended with a

thermostable DNA polymerase (AmpliTaq, Perkin-Elmer) by incubating for 5 minutes at room temperature and then two hours at 70°C. The template was ligated with T4 DNA ligase for 1 hour at 37°C, and then ligated DNA was isolated by ethanol

5   precipitation (Step 3). Before addition of a crosslinking group, the DNA protecting group must be removed (Step 4). This was accomplished by treatment of the oligonucleotide (4.5 μL) with 1.4 M β-mercaptoethanol either at 45°C for 4 hours or at room temperature for 12 hours in 20 μL 20 mM Tris-HCl, pH 8.2.

10   The deprotected DNA was purified to remove the β-mercaptoethanol by ethanol precipitation, and the DNA was dissolved in 20 μL water. A photocrosslinking group was added to the DNA by reaction with the alkylating agent azidophenacyl bromide, and the DNA was again purified for transcription and crosslinking

15   (Step 5).

Photocrosslinking

       For UV crosslinking, the reactions are split in half; a control aliquot is kept in the dark at room temperature while the other aliquot is irradiated at room temperature for 2

20   minutes in a polystyrene tube 1.5 cm from a 302 nm light source (Spectroline model XX-15B, 1800 μW/cm² at 15 cm). After irradiation, DTT (to 60 mM) is added to all samples, and they are left in the dark for at least 5 minutes. After irradiation, but before protein gel electrophoresis, aliquots from the

25   irradiated chase reactions are treated with nuclease to digest the DNA to a small radioactive fragment (Step 6). The nuclease treatment results in digestion of the DNA to small radioactive fragments. This leaves a small radioactive piece of DNA on proteins which have been crosslinked to the DNA during

30   irradiation. The control and irradiated samples are then analyzed by gel electrophoresis and autoradiography.

Electrophoretic Analysis of Crosslinked Proteins

       For identification of crosslinked proteins, samples are mixed with an equal volume of 60 mM TrisCl (pH 8.0), 60 mM DTT,

35   3.4% (w/v) SDS, 17% (v/v) glycerol, 0.02% (w/v) bromophenol blue, 0.02% (w/v) xylene cyanol, heated for 3 minutes at 94°C,

and then analyzed on 12 cm x 0.75 cm 10% SDS/polyacrylamide gels
with a 4% stacking gel (acrylamide/methylene bisacrylamide =
27/1). Proteins are electrophoretically blotted onto MSI
NitroBind Nitrocellulose membrane (Westboro, MA) in 25 mM Tris,
192 mM glycine, 20% (v/v) methanol (Figure 6A). The membrane is
then silver-stained in 2% (w/v) sodium citrate, 0.8% (w/v)
ferrous sulfate, 0.1% (w/v) silver nitrate for 5 minutes and
dried. Autoradiography of the membranes is at -80°C with
Amersham Hyperfilm-MP and a Cronex Lightning Plus intensifying
screen.

5-S-DNP-dU can be used for site-specific modification of
DNA at internal and terminal positions with DNA, and molecular
probes can be placed at variable distances from the DNA
backbone. In addition, the commercial availability of
antibodies to the DNP group allows the direct use of the DNP-
labeled oligonucleotides as immunodetection probes.[15]

Example 4: Utility: Assays for Diagnosis of Bladder Cancer
Bladder Cancer strikes over 50,000 Americans every year,
killing more than 11,000 (23). Currently used means to diagnose
bladder cancer usually rely on immunocytochemistry, nuclear DNA
content and cytology. Extensive use of highly trained labor
that these tests consume make them fairly expensive to
administer. In addition, while being fairly specific for the
diagnosis of bladder cancer, these tests lack sensitivity.
Furthermore, they are poor prognosticators of cancer
progression. Tests or biopsies with better sensitivity and
prognostic value are invasive, even more expensive and poorly
tolerated by patients. These facts result in many patients
being under-diagnosed or under-treated for their illnesses.
An ideal test would be sensitive, specific and prognostic
for bladder cancer, less invasive, require fewer highly skilled
scientists to administer, better tolerated by patients and less
costly. Such an assay would be expected to decrease mortality
and morbidity due to this disease. Such a test would also be
expected to represent a significant commercial possibility since
it would not only replace many of the existing tests in the
marketplace but would enable an increase in the number of tests

48

performed.

One of the methods currently used for detection of mRNA levels in a cell utilizes an RNase protection assay (Hershey and krause, 1989) which is commercially available.   In an RNase
5   protection assay, a labeled RNA probe of discrete length is constructed that includes two domains (Figure 6).   One domain compliments, and therefore hybridizes to, an mRNA species whose abundance one wishes to detect and quantify. This domain forms a double stranded RNA molecule with the intended mRNA target.
10   The other domain is made up of sequence that does not hybridize to the intended target.   In an RNase protection assay, total cell RNA (or total cellular mRNA) is incubated with an excess of an RNA probe under conditions that favor specific hybridization of the probe with the mRNA species to be detected and
15   quantified. After the hybridization step, the RNA is digested with a mixture of RNases. All single stranded RNA is hydrolyzed down to nucleotides or short oligonucleotides.   Only double stranded RNA is protected from this digestion.   All of the unhybridized probe is degraded.   This includes the domains of
20   the hybridizing probes that do not complement their targets.

The digested RNAs are separated according to size by gel electrophoresis and then visualized by autoradiography.   A positive signal is a labeled RNA molecule that is the correct length to be that portion of the probe that had hybridized to
25   the intended target. This shortened form of the original probe can not be created by other means. This is why false positives are not a technical concern. Also, with an excess of probe and under the appropriate hybridization conditions, nearly all of the intended target is hybridized to the nucleic acid probe
30   (NAP). This makes the intensity of the signal proportional to the original concentration of the target mRNA species.

The advantages of an RNase protection assay are that it is quantitative, reproducible and completely free of false positives derived from technical difficulties.   This last
35   feature makes such an assay format appealing as a technological platform for diagnostic determinations of patient derived materials. The disadvantages of an RNase protection assay are that the probes are very tedious to make. They possess very

49

short half lives because of the radioactive isotopes (^{32}P) used to produce the signal. In addition, RNA is inherently less stable than DNA, and some RNases can be extremely difficult to inactivate.

5   The disadvantages of RNase protection are overcome by the envisioned NAP Nuclease Protection Assays (NAP-protection assays) employing the derivatizable nucleotide analogs herein. The NAPs can be synthesized in large quantities, carefully evaluated (quality control assays), and then stored for future
10   use. There is an economy of scale that this affords which overcomes the difficulty of probe construction and synthesis. These NAPs will utilize non-isotopic means of detection, permitting the NAPs to have extended shelf lives. Additionally, the NAPs for different target mRNAs can be modified with
15   different reporter groups. This will permit the examination of the expression of several genes simultaneously. This assay can be modified to accept automation to reduce run times.

Use of oligonucleotides or NAP probes which contain modified nucleotide analogs has been previously reported (21-23,
20   26-29). The chemical approach to preparation of site-specifically modified DNA involves incorporation of a modified nucleotide during the automated synthesis of an oligonucleotide. The modified oligonucleotide can then be incorporated into double-stranded DNA by PCR amplification or primer extension of
25   a single-stranded DNA. The use of various modified phosphoramidites as masked synthons for preparing derivatized oligonucleotides at both terminal and internal sites has been reported (Xu, Zheng and Swann, 1992; Maurizi and Ginsburg, 1986). These masked synthons allow incorporation of protected
30   functional groups such as amines (Nagamachi, et al. 1974), carboxylic acids, thiols (Goodwin and Glick, 1993) and thiocarbonyls. Once incorporated the analogs are deprotected and modified post-synthetically. Convertible nucleoside phosphoramidites, mononomers containing leaving groups, have
35   also been used to incorporate crosslinking and photocrosslinking groups by post-synthetic substitution of the leaving group. Several companies sell reagents for the addition of molecular tags to either of the two ends of a nucleic acid (5′ or 3′).

Some provide reagents which allow the incorporation of more than
one analog at internal positions in the NAP, to allow an
increase in the amount of signal produced by the NAP, there is
a resultant perturbation of nucleic acid structure or decrease
5   in hybridization function (interaction with mRNA) of the NAP.
Since the assay is based on specific NAP hybridization, with the
current methods available, increasing the signal in a NAP must
be balanced against the loss of hybridization capacity.
However, the goal is always to obtain the strongest signal
10   possible, because this not only decreases data acquisition time,
but increases the level of sensitivity of a given assay. This
is clearly an important methodology, for which a commercial
market already exists.

Preferred versions of the nucleotide analogs contemplated
15   herein have the following characteristics which make them
preferable to other currently available nucleotide analogs for
this assay:

1.   The nucleotide modification is on a position of the
base which is not required for normal RNA and DNA
20   basepairing, therefore specific hybridization is still
possible.

2.   The reactive group added to the nucleotide is not
normally found in DNA or mRNA (S group), therefore
reaction of the probe modified oligonucleotides or
25   NAPs with thiol-reactive alkylating agents gives
virtually thiol-specific modification of the NAP.
Analogs with reactive amino groups, which are modified
by alkylating agents targeting amines, are less
specifically modified. This is due to the presence of
30   exocyclic amine groups in the normal nucleotides of
both DNA and RNA, and these can be modified by the
alkylating reagents as well.

3.   The reactivity of the aryl thiol group on 5-SH-UTP to
alkylating agents is considerably higher than that of
35   alkyl thio ether derivatives (pK_A=5.5 vs 8,).

4.   The linker attached to the nucleotide base contains
all single-bonds (vs alkyl or alkynyl), thereby
allowing free rotation of the probe in the region

SUBSTITUTE SHEET (RULE 26)

close to the nucleic acid backbone. This is advantageous because the flexibility of the linker are allows the nucleotide analogs to adopt to different enzyme active sites (in DNA and RNA polymerases),

5   making them generally good substrates for enzymatic incorporation into nucleic acids.

   5.   As assessed by several functional assays, analogs modified through the aryl thio group of 5-S-U derivatives do not disturb normal RNA secondary or

10   tertiary structures.

Unlike other types of cancer, the primary tumor in bladder cancer sheds or exfoliates relatively large numbers of cells into the lumen of the bladder. These cells can be easily collected in urine or from bladder washes. In a void urine from

15   a patient with a bladder tumor, there will be very few to about a million cancer cells. Bladder washes yield more cells than void urines. Currently, the company UroCor, Inc. and other firms provide services to diagnose disease. Bladder cancer cells collected in urine are examined for morphology, DNA

20   content and for protein biomarkers as detected immunochemically. While these tests are not as informative as desired, it is clear that cancer cells can be collected from patient urines and that these cells are relatively intact with respect to morphology, nucleic acid and protein content, and expression. RNA has not

25   been directly examined in these cells, but it is reasonable to expect that the RNA content of these cells will also be intact relative to the needs of the envisioned NAP-protection assay.

A NAP-protection assay can be designed to determine the relative abundance of mRNAs for four surrogate biomarkers of

30   bladder cancer and/or its progression relative to the abundance of a mRNA encoding a "house keeping" gene that is not differentially regulated. Changes in the abundance of the surrogate biomarkers relative to the undifferentially expressed gene will be informative in both diagnosing and evaluating the

35   state of a patient's bladder cancer. In one embodiment, the final assay uses five probes, each specific for a different gene. One of the probes will hybridize to an appropriate "house keeping" gene. The hybridization signal from this probe will

52

act as one of the positive controls in the assay. Of the remaining four mRNA species, targets will be selected such that in cancerous cells two of the mRNAs are increased in abundance and two are decreased relative to nontransformed cells. In this
5   format, each regulatory pathway acts as an internal control for the other.

The following assays can be carried out to probe for mRNA levels of the human autocrine motility factor receptor, which is differentially expressed in bladder cancer cells. The DNA
10   sequence (SEQ. ID. NO:4) for this gene is shown in Figure 7.

Ribonuclease NAP Protection Assay

One ribonuclease assay is based upon the currently utilized radioisotopic RNase protection assay, described above. However, the radioactively labeled nucleotide substrates are replaced by
15   5-S-R-XTP nucleotides, where X is U, dU, C, or dC and R is one of several different reporter groups. The NAPs for the assay are ssRNA, which are synthesized enzymatically utilizing a bacterial or phage RNA polymerase.

In this assay, RNA NAPs can be prepared in which the RNA is
20   labeled with many fluorescent, colormetric, chemiluminescent, or antigenic reporter groups. One such antigenic reporter group is the dinitrophenyl group (DNP), and the analog for these experiments is proposed below. The assay is the same as the current method previously described, but the protected RNA
25   products are detected by either monitoring fluorescent emission, probing with antibodies, or monitoring chemiluminescence. In this assay, final detection still requires the analysis of the NAP protection assay by gel electrophoretic separation of the products. However, the probes can be synthesized in large
30   quantities and stored, so this assay still has a major advantage over the existing methodologies.

Deoxyribonuclease NAP Protection Assay

An alternative nuclease protection assay can utilize DNA NAPs, similarly modified with multiple reporter groups. One
35   compelling reason to consider DNA NAPs is that they are both chemically and biologically more stable than the comparable RNA

53

NAPs. This makes them better prospects for synthesis in large quantities and for long storage. Despite this inherent stability relative to RNA, ssDNA can be easily degraded by single-strand specific DNases, in the manner as utilized for
5   ssRNA digestion. In addition, with DNA NAPs, one can synthesize large quantities of the DNA probe by utilizing the DNA amplification method, PCR (polymerase chain reaction).

1)   DNA NAPs Prepared with Analog-Tagged Oligonucleotide
Primers
10   Already developed for use in such a DNA NAP-protection assay is the analog 5-DNP-dU phosphoramidite. This analog can be incorporated chemically into ssDNA oligonucleotides utilizing automated synthesis. Therefore oligonucleotide primers for the PCR reaction can be made which contain multiple analogs, with
15 nearly every "T" in the oligonucleotide synthesis and isolation, and we have shown that it is also retained during the thermal cycling reaction of PCR. Therefore, one can prepare DNA NAPs which contain several reactive thiol groups in the 5' region of the NAP, and these can then be deprotected and modified with a
20 variety of alkylating agents. One can therefore prepare probes for several different mRNA targets, each with a different reporter group, and differentially quantify the levels of each in one reaction. This aspect of the method becomes important for the progression of the assay to automation.

25   2)   DNA NAPs Prepared with Analog-Tagged Deoxynucleoside
Triphosphates
DNA NAPs can be synthesized by incorporating several (e.g., five to ten) 5-S-DNP-dUs into the 5'-tagged oligonucleotide, and then synthesizing full length DNA NAP by PCR with unmodified
30 dNTPs, e.g., 5-S-DNP-dUTP.

Automation of NAP Protection Assays
Automation of either the RNA NAP Protection Assay or the DNA NAP Protection Assay requires that the electrophoresis step be eliminated. The separation of the individual protected mRNA
35 species by electrophoresis is one of the more time-consuming

SUBSTITUTE SHEET (RULE 26)

54

steps of the current assay, and makes this method intractable to automation. In order to eliminate this step, one must devise a method for the detection and relative quantitation of more than one mRNA species in a single reaction. Currently, the best

5   approach for this involves the use of a variety of fluorescent tags, each with different excitation and emission properties. An increasing number of alkylating reagents tagged with fluorescent groups such as fluorescein or rhodamine, to name only two of many, are now commercially available. Companies are

10   designing fluorescent detectors which can differentiate among the different signals produced by these fluorescent dyes. This approach is modeled upon the automated assays currently used for nonisotopic DNA sequencing. The final choice of fluorescent dye combinations used for these assays are determined empirically

15   during the course of the first year of this grant. Alternatively, one can label different NAPs with different antigenic groups, so that RNAs could be distinguished with antibodies, a method already utilized in many automated assays.

Deoxyribonuclease NAP Protection Assay

20   For the DNA NAP assay, one can incorporate commonly used methods for the immobilization of DNA onto resins or microtiter plates. If, for example, a 96-well microtiter plate were coated with streptavidin, then DNA tagged at the 5' end with biotin could be efficiently immobilized. In this way, the DNA NAP and

25   any RNA to which it hybridizes, could be immobilized on a solid matrix, while other components of the hybridization mixture could be washed away.

The following approach can be followed for establishment of an automated DNase NAP Protection Assay.

30   1)   DNA oligonucleotides can be synthesized which are tagged at the 5' end with one of many groups that would allow immobilization of the DNA to a solid matrix (for example biotinylated nucleotides and streptavidin-agarose resin). The DNA NAPs can be multiply-labeled by either:

35   a)   incorporating several (5-10) 5-S-DNP-dUs into the 5'-tagged oligonucleotide, and then synthesizing

SUBSTITUTE SHEET (RULE 26)

full length DNA NAP by PCR with unmodified dNTPs.

b)   incorporating multiple (50-100) 5-S-DNP-dUs into the
        body of the DNA product produced by PCR by using 5-
        S-DNP-dUTP as a substrate.

5   2)   The 5' end-labeled, internally reporter-labeled DNA
NAPs can be deprotected and modified with one of the reporter
groups chosen in the studies described above.

        3)   Reaction mixtures can be prepared containing multiple
mRNA targets and one or more reporter-tagged NAPs.

10   4)   After hybridization, reactions mixtures can be treated
with nucleases specific for single-stranded nucleic acids (such
as S1 or Mung Bean Nucleases), to produce the protected mRNA
species.

        5)   Reactions mixtures can be added to microtiter plates
15   coated with Streptavidin (for 5'-biotin labeled DNA) or other
materials which are specific for binding the groups commonly
used to tag DNA probes.   Since the 5' end of the DNA is
protected during the NAP-protection assay (see Fig. 1), and NAP
which does not   hybridize are degraded by nuclease, the only
20   material which   should bind to the solid matrix is the DNA-
mRNA hybrid.

        6)   Microtiter   plates   can   be   washed   to   remove
unincorporated   nucleotides, thereby removing essentially all
background signal.

25   7)   Plates can be analyzed with a fluorescent microtiter
plate reader to quantify the relative amounts of mRNAs produced
for different NAPs.

REFERENCES

Agrawal, S., ed. <u>Methods in Molecular Biology</u>, ed. Walker, J.M.,
    Vol. 20, Humana Press: Totowa, New Jersey (1993).

Barone, A.D., Tang, J.-Y. and Caruthers, M.H. (1984) *In situ*
    Activiation of bis-Dialkylaminophosphines - A New Method
    for Synthesizing Deoxyoligonucleotides on Polymer Supports
    Nucleic Acids Res. 12, 4051-4061.

Bartholomew, B., Braun. F.R., Kassavetis, G.A. and Geiduschek,
    E.P. (1994) Probing Close DNA Contacts of RNA Polymerase
    III Transcription Complexes with the Photoactive Nucleoside
    4-Thiodeoxythymidine *J. Biol. Chem. 269*, 18090-18095.

Bartholomew, B., Kassavetis. G.A., Braun, B.R. and Geiduschek,
    E.P. (1990) The Subunit Structure of *Saccharomyces
    cerevisiae* Transcription Factor IIIC Probed with a Novel
    Photocrosslinking Reagent *EMBO J. 9*, 2197-2205.

Beaucage, S.L. and Iyer, R.P. (1993) The Functionalization of
    Oligonucleotides Via Phosphoramidite Derivatives
    Tetrahedron 49, 1925-1963.

Bergstrom, D.E., Beal, P., Jenson, J. and Lin, X. (1991)
    Palladium-mediated Synthesis of C-5 Pyrimidine Nucleoside
    Thioethers from Disulfides and Mercurinucleosides *J. Org.
    Chem. 56*, 5598-5602.

Bleasdale, C., Ellwood, S.B. and Golding, B.T. (1990) 4,4'-
    Dimethoxytrityl and 4-Monomethoxytrityl Tetrafluoroborate:
    Convenient Reagents for the Protection of Primary Alcohols
    Including Sugars *J. Chem. Soc. Perkin Trans 1* 803-804.

Bradley, D.H. and Hanna, M.M. (1992) Synthesis and Utility of 5-Thiocyanato Dexoyuridine and Uridine Phosphoramidites as Masked Synthons *Tetrahedron Lett.* 33, 6223-6226.

Brown, H.C., Narasimhan, S. and Choi, Y.M. (1982) Selective Reduction. 30. Effect of Cation and Solvent on the Reactivity of Saline Borohydrides for Reduction of Carboxylic Esters. Improved Procedures for the Conversion of Esters to Alcohols by Metal Borohydrides *J. Org. Chem.* 47, 4702-4708.

Chaudhary, S.K. and Hernandez, O. (1979) A Simplified Procedure for the Preparation of Triphenylmethyleters *Tetrahedron Lett.* 95-98.

Christopherson, M.S. and Broom, A.D. (1991) Synthesis of Oligonucleotides Containing 2'-Deoxy-6-thioguanosine at a Predetermined Site *Nucleic Acids Res.* 19, 5719-5724.

Clegg, R.M., Murchie, A.I.H., Zechel, A., Carlberg, C., Diekmann, S. and Lilley, D.M.J. (1992) Fluorescence Resonance Energy Transfer Analysis of the Structure of the Four-Way DNA Junction *Biochemistry* 31, 4846-4856.

Clivio, P., Fourrey, J.-L., Gasche, J., Audic, A., Favre, A., Perrin, C.; and Woisard, A. (1992) Synthesis and Purification of Oligonucleotides Containing Sulfur Substituted Nucleobase: 4-Thiouracil, 4-Thiothymine and 6-Mercaptopurine *Tetrahedron Lett.* 33, 65-68.

Coleman, R.S. and Kesicki, E.A. (1994) Synthesis and Postsynthetic Modification of Oligodeoxynucleotides Containing 4-Thio-2'-deoxyuridine (d^{s4}U) *J. Am. Chem. Soc.* 116, 11636-11642.

Connolly, B.A. and Newman, P.C. (1989) Synthesis and Properties of Oligonucleotides Containing 4-Thiothymine, 5-Methy-2-pyrimidinone-1-β-D-(2'-deoxyriboside) and 2-Thiothymidine *Nucleic Acids Res.* 17, 4957-4974.

Cosstick, R. and Douglas, M.E. (1991) Synthesis of a Dinucleoside Monophosphate Analogue Containing 6-N-(2-Aminoethyl)-2'-Deoxyadenosine. A Novel Approach to Sequence Specific Crosslinking in Synthetic Oligonucleotides *J. Chem. Soc. Perkin Trans. 1* 1035-1040.

Dissinger, S. and Hanna, M.M., (1990) <u>J. Biol. Chem.</u>, 265:7662.

SUBSTITUTE SHEET (RULE 26)

58

Flaschka, H.A., Barnard, A.J. and Sturrock, P.E. (1969) *Quantitative Analytical Chemistry, Short Introduction to Practice*, Barnes & Noble, INC., New York.

Gait, M.J., ed. Practical Approach Series, ed. Rickwood, D. and Harnes, B.D., IRL Press: Oxford (1984).

Gao, H., Fathi, R., Gaffney, B.L., Goswami, B., Kung, P.-P., Rhee, Y., Jin, R. and Jones, R.A. (1992) 6-O-(Pentatluorophenyl)-2'-deoxyguanosine: A Versatile Synthon for Nucleoside and Oligonucleotide Synthesis *J. Org. Chem.* 57, 6954-6959.

Glick, G.D. (1991) Synthesis of a Conformationally Restricted DNA Hairpin *J. Org. Chem. 56*, 6746-6747.

Gibson, K.J. and Benkovic, S.J. (1987) Nucleic Acid Res., 15, No. 16:6455.

Goodwin, J.T. and Glick, G.D. (1993) Incorporation of Alkylthiol Chains at C-5 of Deoxyuridine *Tetrahedron Lett. 34,* 5549-5552.

Goodwin, J.T., Osborne, S.E., Swanson, P.C. and Glick, G.D. (1994) Synthesis of a Disulfide Cross-linked DNA Triple Helix *Tetrahedron Lett. 35*, 4527-4530.

Greene, T.W. (1981) *Protective Groups in Organic Synthesis*, John Wiley & Sons, New York.

Grzybowski, J., Will, D.W., Randall, R.E., Smith, C.A. and Brown, T. (1993) Synthesis and Antibody-mediated Detection of Oligonucleotides Containing Multiple 2,4-Dinitrophenyl Reporter Groups *Nucleic Acids Res. 21*, 1705-1712.

Hanna, M.M. (1989) Methods in Enzymology, 180:383

Hanna, M.M. (1993) Photocrosslinking Analysis of Protein-RNA Interactions in *E. coli* Transcription Complexes *Cellular Molecular Biology Research 39*, 393-399.

Hanna, M.M., Dissinger, S., Williams, B.D. and Colston, J.E. (1989) Synthesis and Characterization of 5-[4-(Azidophenacyl)thio]uridine 5'-Triphosphate, a Cleavable Photo-Cross-Linking Nucleotide Analogue *Biochemistry 28*, 5814-5820.

He, B. and Hanna, M.M., (1995) Nucleic Acids Res., 23.

Hershey, A.D. and Krause, J.E., (1989) Science, 247:960.

Jadhav, V., Barawkar, D., Natu, A., and Ganesh, K. *Nucleosides*

SUBSTITUTE SHEET (RULE 26)

59

and Nucleotides, (1997) 16, 107-114.

Kalman, T.I. and Bardos, T.J. (1967) The Autoxidation of 5-Mercaptouracil and 5-Mercaptodeoxyuridine *J. Am. Chem. Soc.* *89*, 1171-1175.

Keller, G.H., and Manak, M.M., (1989) <u>DNA Probes</u>, New York, N.Y.: Stockton Press.

Kini, G.D., Robins, R.K. and Avery, T.L. (1989) Synthesis and Antitumor Activity of Ribavirin Imidates. A New Facile Synthesis of Ribavirin Amidine (1-β-D-Ribofuranosyl-1,2,4-triazone-3-carboxamidine Hydrochloride) *J. Med. Chem. 32*, 1447-1449.

Kricka, L.J. (1992) ed. <u>Nonisotopic DNA Probe Techniques</u>, Academic Press, Inc.: San Diego, CA.

Kuimelis, R.G. and Nambiar, K.P. (1994) Synthesis of Oligodeoxynucleotides Containing 2-Thiopyrimidine Residues-A New Protection Scheme *Nucleic Acids Res. 22*, 1429-1436.

Lakshman, M.K., Sayer, J.M., Yagi, H. and Jerina, D.M. (1992) Synthesis and Duplex-Forming Properties of a Nonanucleotide Containing an N^6-Deoxyadenosine Adduct of a Bay-Region Diol Epoxide *J. Org. Chem. 57*, 4585-4590.

Maurizi, M.R. and Ginsburg, A. (1986) <u>Biochemistry</u>, 25:131.

Mergny, J.-L., Boutorine, A.A., Garestier, T., Belloc, F., R ougee, M., Bulychev, N.V., Koshkin, A.A., Bourson, J., Lebedev, A.V., Valeur, B., Thuong, N.T. and Helene, C. (1994) Fluorescence energy transfer as a probe for nucleic acid structures and sequences *Nucleic Acids Res. 22*, 920,928.

Nagamachi, T., Fourrey, J.-L., Torrence, P.F., Waters, J.A. and Witkop, B. (1974) Synthesis, Chemistry, and Biological Activity of 5-Thiocyanatopyrimidine Nucleosides as Potential Masked Thiols *J. Med. Chem. 17*, 403-406.

Scremin, C.L., Zhou, L., Srinivasachar, K. and Beaucage, S.L. (1994) Stepwise Regeneration and Recovery of Deoxyribonucleoside Phosphoramidite Monomers during Solid-Phase Oligonucleotide Synthesis *J. Org. Chem. 59*, 1963-1966.

Shaltiel, S. (1967) Thiolysis of Some Dinitrophenyl Derivatives of Amino Acids *Biochem. Biophys. Res. Comm. 29*, 178-183.

SUBSTITUTE SHEET (RULE 26)

60

Sheardy, R.D. and Seeman, N.C. (1986) The Synthesis of a Deoxyoligonucleotide Incorporating 5-Iododeoxyruidine *J. Org. Chem.* *51*, 4301-4303.

Sinha, N.D., Biernat, J., McManus, J., and Koster, H. (1984) Polymer Support Oligonucleotide Synthesis XVIII: Use of β-cyanoethyl-*N,N*-dialkylamino-/*N*-morpholino Phosphoramidite of Deoxynucleosides for the Synthesis of DNA Fragments Simplifying Deprotection and Isolation of the Final Product *Nucleic Acids Res.* *12*, 4539-4557.

Soai, K. and Ookawa, A. (1986) Mixed Solvents Containing Methanol as Useful Reaction Media for Unique Chemoselective Reductions with Lithium Borohydride *J. Org. Chem.* *51*, 4000-4005.

Torrence, P.F., Waters, J.A. and Witkop, B. (1968) 2'Deoxy-5'(thiocyanato)uridine [1-(2-deoxy-β-D-erythropentofuranosyl)-5-(thiocyanato)uracil] *Synthetic Procedures in Nucleic Acid Chemistry*, (W.W. Zorbach and R.S. Tipson, Ed.) pp 367-370, Interscience, New York.

Trivedi, B.K. and Bruns, R.F. (1989) C2.N^6-Disubstituted Adenosines: Synthesis and Structure-Activity Relationships *J. Med. Chem.* *32*, 1667-1673.

Waters, T.R. and Connolly B.A. (1992) Straightforward Synthesis of 6-Thiodeoxyguanosine and Its Incorporation into Oligodeoxynucleotides *Nucleosides Nucleotides* *11*, 985-998.

Xu, Y.-Z., Zheng, Q. and Swann, P.F. (1992a) Synthesis and Duplex Stability of Oligodeoxynucleotides Containing 6-Mercaptopurine *Tetrahedron Lett.* *33*, 5837-5840.

Xu, Y.-Z., Zheng, Q. and Swann, P.F. (1992b) Synthesis by Post-Synthetic Substitution of Oligomers Containing Guanine Modified at the 6-Position with S-, N-, O-derivatives *Tetrahedron* *48*, 1729-1740.

Xu, Y.-Z., Zheng, Q. and Swann, P.F. (1992c) Synthesis of DNA Containing Modified Bases by Postsynthetic Substitution. Synthesis of Oligomers Containing 4-Substituted Thymine: O^4-Alkylthymine, 5-Methylcytosine, N^4-(Dimethylamino)-5-methylcytosine, and 4-Thiothymine *J. Org. Chem.* *57*, 3839-3845.

Yang, X.J., Hart, C.M., Grayhack, E.J. and Roberts, J.W. (1987)

61

Transcription Antitermination by Phage Lambda Gene Q
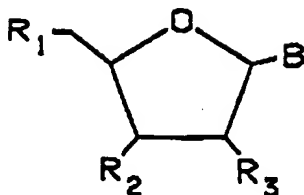Protein Requires a DNA Segment Spanning the RNA Start Site
Genes Dev 1, 217-226.

Zhang, Y. and Hanna, M.M. (1995) Expression and Functional
Characterization of *Escherichia coli* NusA and Lambda Q as
Glutathione-S-Transferase Fusion Proteins *Protein
Expression and Purification 6*, 625-631.

62

Changes may be made in the compounds described herein or in the steps or the sequence of steps of the methods of synthesis or use of the compounds described herein without departing from the spirit and scope of the invention as defined in the following claims.

63

What is claimed is:

1.   A nucleotide analog of the formula:

wherein

R₁ is -H, -OH, a mono, di, or triphosphate group, or -OR₄;

R₂ is -H, -OH, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, -OR₄, or a solid support bonded via an O at the 3' position;

R₃ is -H, -OH, a mono, di, or triphosphate group, or -OR₄;

R₄ is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at the 5 position on said base that is not involved in Watson-Crick base pairing or does not disrupt normal Watson-Crick base pairing, said protected thiol group being stable under conditions of chemical nucleic acid synthesis and/or conditions of enyzmatic nucleic synthesis and being convertible to a reactive thiol after said synthesis.

2.   The nucleotide analog of Claim 1 wherein said base is selected from the group consisting of cytosine and uracil.

3.   The nucleotide analog of Claim 1 wherein said base compromises an additional protecting group on a reactive moiety of said base.
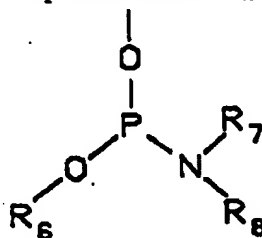
4.   The nucleotide analog of Claim 3 wherein said base is selected from the group consisting of N⁴-anisoyl cytosine, N⁴-

SUBSTITUTE SHEET (RULE 26)

64

benzoyl cytosine, N^4-isobutyryl cytosine and N^4-acetyl cytosine.

5.   The nucleotide analog of Claim 1 wherein said protected thiol group on said base is selected from the group consisting of thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R_{10}-DNP), and alkyldisulfide (-S-S-R_{10}) wherein R_{10} is a lower alkyl.

6.   The nucleotide analog of Claim 1, wherein the protecting group of R_1 is selected from the group consisting of p-(dimethoxytrityl), p-(monomethoxytrityl), fluorenylmethyl-oxycarbonyl, levuloyl and 9-phenylxanthene-9-yl.

7.   The nucleotide analog of Claim 1, wherein protecting group of R_3 is selected from the group consisting of 1-(2-chloro-4-methylphenyl)-4-methoxy-4-piperidinium, 2'-acetal, o-nitrobenzyl, tert-butyldimethyl silyl, tetrahydrofuranyl and 4-methoxytetrahydropyranyl.

8.   The nucleotide analog of Claim 1 wherein R_1 is -OR_4 wherein R_4 is a protecting group and R_2 is a phosphoramidite group.

9.   The nucleotide analog of Claim 8 wherein said base is cytosine, or uracil.

10.   The nucleotide analog of Claim 9 wherein said base comprises an additional protecting group on a reactive moiety of said base.

11.   The nucleotide analog of Claim 12 wherein said base is selected from the group consisting of N^4-anisoyl cytosine, N^4-benzoyl cytosine, N^4-isobutyryl cytosine and N^4-acetyl cytosine.

12.   The nucleotide analog of Claim 8 wherein said protected thiol group of said base is selected from the group consisting of thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R_{10}-DNP), and alkyldisulfide (-S-S-R_{10}) wherein R_{10} is a lower alkyl.

SUBSTITUTE SHEET (RULE 26)

65

13. The nucleotide analog of Claim 8 wherein the protecting group of R_1 is selected from the group consisting of p-(dimethoxytrityl), p-(monomethoxytrityl), fluorenylmethyl-oxycarbonyl, levuloyl, and 9-phenylxanthene-9-yl.
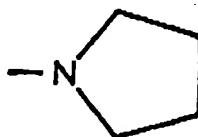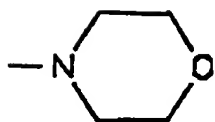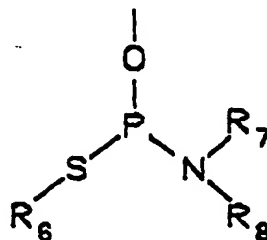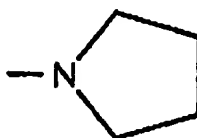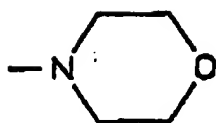
14. The nucleotide analog of Claim 8, wherein the protecting group of R_3 is selected from the group consisting of 1-(2-chloro-4-methylphenyl)-4-methoxy-4-piperidinyl,2'-acetal, o-nitrobenzyl, tert-butyldimethyl silyl, tetrahydrofuranyl and 4-methoxytetrahydropyranyl.

15. The nucleotide analog of Claim 1 or 8, wherein said phosphoramidite of R_2 is represented by the formula:

R_6 is a lower alkyl, cyanoethyl or a substituted lower alkyl; and

R_7 and R_8 are independently lower alkyls, or when taken together with the nitrogen to which they are attached comprise one of the groups:

, or

16. The nucleotide analog of Claim 1 wherein R_1 is $-OR_4$ wherein R_4 is a protecting group and R_2 is a phosphorothioamidite group.

66

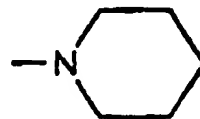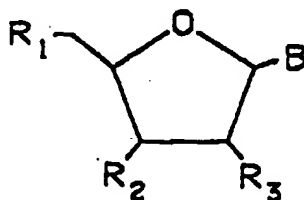17. The nucleotide analog of Claim 1 or 16, wherein said R₂ is a phosphorothioamidite represented by the formula:

wherein

R₆ is a lower alkyl, cyanoethyl or a substituted lower alkyl; and

R₇ and R₈ are independently lower alkyls, or when taken together with the nitrogen to which they are attached comprise one of the groups:

18. The nucleotide analog of Claim 1 wherein R₂ is an O-substituted monophosphate group selected from the group consisting of -2-chlorophenyl monophosphate, O-2,5-dichlorophenyl monophosphate, O-2,2,2-trichloroethyl monophosphate and the N oxide of 4-methoxypyridine-2-methylene monophosphate.

19. A nucleotide analog of the formula:

wherein

R₁ is -H, -OH, a mono, di, or triphosphate group, or -OR₄;

R₂ is a phosphoramidite group;

SUBSTITUTE SHEET (RULE 26)
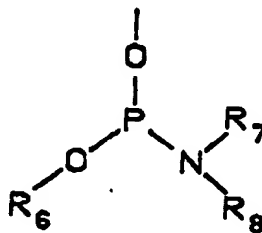
R₃ is -H, -OH, a mono, di, or triphosphate group, or -OR₄;

R₄ is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a thiodinitrophenyl group attached at the 5 position on said pyrimidine base.

20.   The nucleotide analog of Claim 19 wherein said base is selected from the group consisting of N⁴-anisoyl cytosine, N⁴-benzoyl cytosine, N⁴-isobutyryl cytosine and N⁴-acetyl cytosine.

21.   The nucleotide analog of Claim 19, wherein R₁ is -OR₄ and R₄ is a protecting group selected from the group consisting of p-(dimethoxytrityl), p-(monomethoxytrityl), fluorenylmethyl-oxycarbonyl, levuloyl, and 9-phenylxanthene-9-yl.
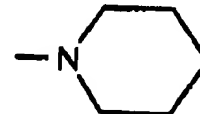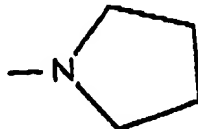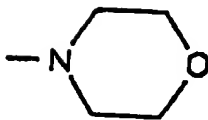
22.   The nucleotide analog of Claim 19 wherein said phosphoramidite of R₂ is represented by the formula:

wherein

R₆ is a lower alkyl, cyanoethyl or substituted lower alkyl; and

R₇ and R₈ are independently lower alkyls, or when taken together with the nitrogen to which they are attached form the groups:
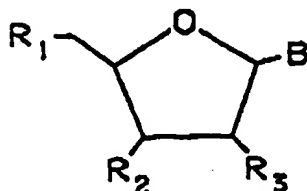

23.   The nucleotide analog of Claim 22 wherein R₆ is 2-cyanoethyl, R₇ is isopropyl and R₈ is isopropyl.

68

24.   The nucleotide analog of Claim 23 wherein said base is uridine or deoxyuridine.

25.   The nucleotide analog of Claim 24 wherein R_1 is $-OR_4$ and R_4 is p-(dimethoxytrityl).

26.   A nucleotide analog of the formula:

wherein:

R_1 is $-H$, $-OH$, a mono, di, or triphosphate group, or $-OR_4$;

R_2 is $-H$, $-OH$, a mono, di, or triphosphate group, a phosphoramidite group, a phosphorothioamidite group, a phosphonate group, an O-substituted monophosphate group, $-OR_4$, or a solid support bonded via an O at the 3' position;

R_3 is $-H$, $-OH$, a mono, di, or triphosphate group, or $-OR_4$;

R_4 is a lower alkyl or a protecting group; and

B is a modified pyrimidine base comprising a protected thiol group attached at the 5 position on said base.

27.   The nucleotide analog of Claim 26 wherein said base is cytosine, or uracil.

28.   The nucleotide analog of Claim 26 wherein said protected thiol group on said base is selected from the group consisting of thiodinitrophenyl (-SDNP), thioalkyldinitrophenyl (-S-R_{10}-DNP), S-phenylacetamidomethyl(-S-$CH_2NHCOCH_2Ph$) wherein R_{10} is a lower alkyl.

29.   The nucleotide analog of Claim 26, wherein R_1 is a monophosphate, a diphosphate, or a triphosphate.

69

30. An oligonucleotide containing the nucleotide analog of any one of Claims 1, 8, 16, 19, or 26.

31. A method of producing the nucleotide analog of any one of Claims 1, 8, 16, 19, or 26 which comprises preparing a thiol-protected nucleoside or nucleotide base wherein said thiol is attached to the 5 position on said base; reacting said nucleoside or nucleotide base under conditions to effect conversion of said base to a phosphoramidite, phosphorothioamidite, phosphonate, O-substituted monophosphate or phosphate nucleotide analog and under conditions which do not destroy the protected thiol; and recovering said analog.

5

32. A method of synthesizing a nucleic acid having an attached functional group which comprises incorporating a thiol-protected nucleotide analog of any one of Claims 1, 8, 16, 19, or 26 into a nucleic acid by a chemical or enzymatic method of nucleic acid synthesis; recovering said nucleic acid containing said analog; deprotecting the analog of said nucleic acid to produce a nucleic acid containing a reactive thiol group; treating the reactive thiol group with a thiol modifying reagent to thereby attach a functional group and produce said nucleic acid with an attached functional group; and recovering said nucleic acid with said attached functional group.

5

10

33. The method of Claim 32 wherein said functional group is selected from the group consisting of a photocrosslinker, a crosslinker, a reporter molecule, a radioisotope, a fluorescent group, a spin label, chemiluminescent or an antigenic group.

34. The method of Claim 33 wherein said photocrosslinker is an aryl azide.

35. The method of Claim 33 wherein said reporter group is biotin, an enzyme, or a fluorescent molecule.

36. The method of Claim 33 wherein said fluorescent group is fluorescein.

SUBSTITUTE SHEET (RULE 26)

37. A method for producing a deoxyuridine analog having a protected thiol group on the 5 position, comprising:

      reacting a diacetyl 5-thiocyanate deoxyuridine to form a 5-dinitrophenylthio deoxyuridine;

      converting said 5-dinitrophenylthio deoxyuridine to a phosphoramidite, phosphorothioamidite, phosphonate, O-substituted monophosphate or phosphate nucleotide analog under conditions which do not destroy the 5-dinitrophenylthio group;

      and recovering said 5-dinitrophenylthio deoxyuridine analog.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9401445 A	20-01-94	NONE	

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category°	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	M. M. HANNA AND K.L. MEYER: "Synthesis and Characterization of a New 5-Thiol-Protected Deoxyuridine Phosphoramidite for Site-Specific Modification of DNA" BIOCONJUGATE CHEMISTRY, vol. 7, no. 4, 1996, pages 401-412, XP002045926 see the whole document ---	1-37
A	J. CHATTOPADHYAYA ET AL.: "Synthesis of C-5 and N-3 Arenesulfenyl Uridines" ACTA CHEM. SCAND. SER. B, vol. 39, 1985, pages 203-12, XP002045927 -----	

1

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6   C07H21/00   C07H19/10   C07H19/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6   C07H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category°	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 01445 A (RES CORP TECHNOLOGIES INC) 20 January 1994 see the whole document ---	1-4, 6-11, 13-18, 26,27, 29-36
X	M.M. HANNA: "Photochemical Cross-Linking Analysis of Protein-Nucleic Acid Interactions in Escherichia coli Transcription Complex from Lambda PR' Promoter" METHODS IN ENZYMOLOGY, vol. 274, 1996, pages 403-418, XP002045925 see the whole document ---	1-37

-/--

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 November 1997	1 6. 12. 97

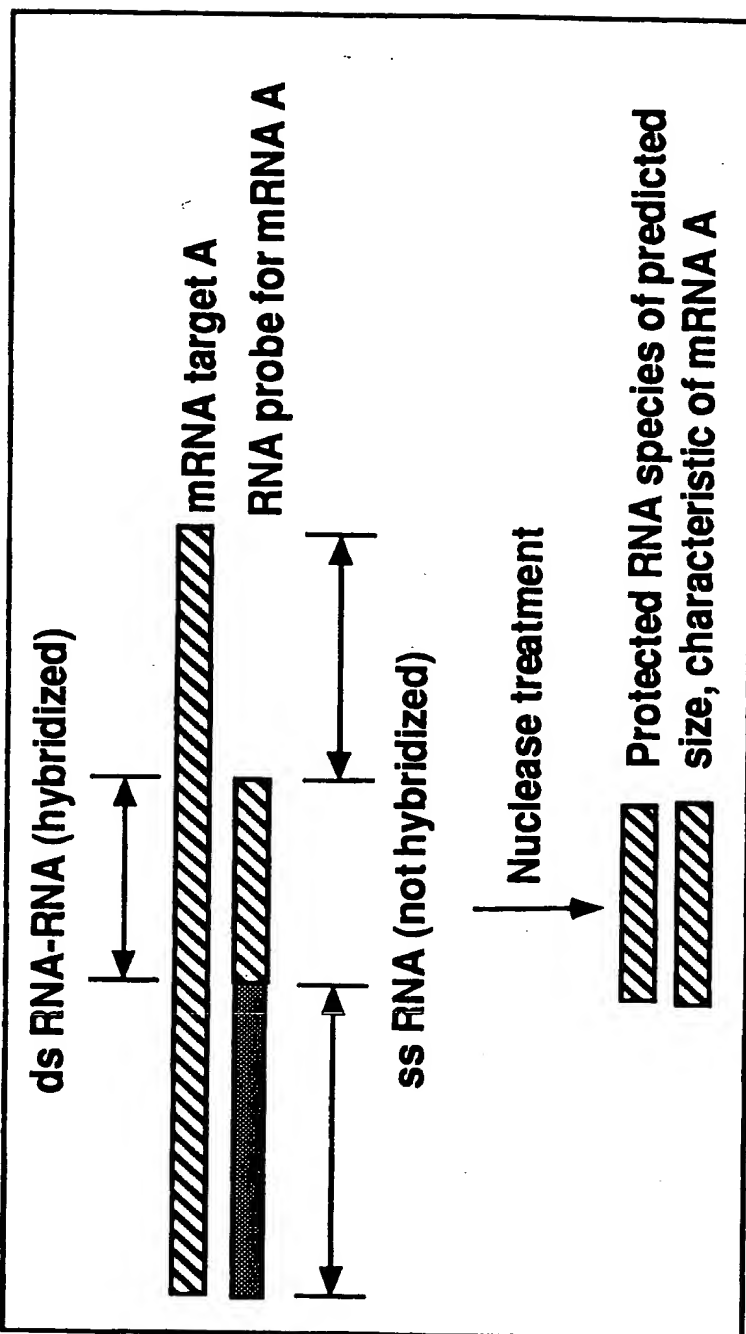| Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bardili, W

1

Form PCT/ISA/210 (second sheet) (July 1992)

page 1 of 2

FIGURE 7: DNA Sequence of Human Autocrine Motility Factor Receptor (AMFR)

```
   1 GAAACTGCCT GTGAAACATCT TTTCCACAACTCCTGTCCTC GTTCCTGCCT AGAACAAGAC
  61 ACCTCCTGTC CAACATGCAG AATGTCTCTT AATATTGCCG ACAATATCGT GTCAGGGAAG
 121 AACATCAAGG AGAGAAACTTG GATGAGAATT TGGTTCCTGA TGCAGCAGCC GAAGGAGACC
 181 TGGCTTAAAC CAACACAATC ACTTCTTCCA TTTCGATGGG TCTCGGATTG CGAGCTGGCT
 241 GCCGAGTTTT TCGGTTGAAG TGATGCACAC CACCAACATT CTTGGCATTA CGCAGGCGAG GCCAGGCAG
 301 CAACTCCCAG CTCAATGCAA TGCCTCATCA GATTCAAGAG ATGTTTCCCC AGGTTCCATA
 361 CCATCTGGTA CTGCCAGGACC TCCAGCTGAC ACGCTCAGTT AAATAACAAC AGTCAATTATT
 421 TTAGAAGGAC GGATTCAACT ACCTTTTCCT ACACAGCCGT CAGATAGCAT CAGACCTGCA
 481 TTGAACAGTC CTGTGGAAAG GCCAAGCAGT GACCAGGAAG AGGGAGAAACTTCTGCTCAG
 541 ACCGAGCCGTG GTTGCCACTCGG ACCTCCAGTCC TCCCCCTCGGAG GAGACGCTCG ACTTCGGGGGA
 601 GGTTGGGAAGTG GAGCCCCAGTTG AGGTTGGAAGA CCTTCGGAGGGCT CGTTGGGGAGCGC GCCTTCCTCCAA
 661 GTCTGCTGAT GAGAGACAGC GCATGCCTGGT CGCAGCCGGTAA GGACGAACTC CCTGCAGCAAG
 721 CTCGCCAAACG TTTCTTGAAC AAAAGTTCTG AAGATGGATCC GGCCTCAGAG AGCTTCCTCC
 781 CCTTGGGAAGG TGCGTCCCTCT GACCGGGTGA CCCTGCGTCG AAGGATGCTG GCTGCCGGGGC
 841 GGAACGGGAGG CTTCAGAAGCA GCAGGACCTC CTAGCGCTCC CTTGCCTTCC TCAGCTGCCT
 901 CCCTGCGCCCT GTGCCGGACT CACTGGAGGA GGCCCTGTCCC AATTCTGCCG CTCCATGGAA
 961 AAGCGGGCTT GACTGCCATTG CCGCCTGTATA AAGGCATGTGG TCTTATAGTG TTTGGACAGC
1021 TGATAAATTT AATCCTTCTT TGTAATACTT TCTATGTGAC ATTTCTCTTC CCCTTAGAAA
1081 CACTTGCAAAT TTTAACTGTA GGTATGATCT CTTCTTGGTTGT TGACTGGACT GCCTTGGGGGTG
1141 GGGGACGGATC AGGAGGAAGT GAGCCCAGTCG CCTGCCTGCA GCAGGGCAGCT TCTACTCCTG
1201 CCCTCATGCAT AGGTCCCACA AATGCAGGGTG TCCCTGAGCAC CACACCCCAGT GGGAAGAGTG
1261 TTGGGGGAGGC GCACAGTGTG AGCCOGCGCCC CACGTGGTGG GGTAACATCT GTTATTCAAAC
1321 TGCTGTGGGTT GTTGTGGAAG CATGTAGACT GTGCCAGAGC CAGACCCACG GGCTCATGCA
1381 CCCCTGAGCA GCCAGGGCCATC TTGGAAAAGG AACTCTTGGT TCGATAOCTTG GAGCAGAGGA
1441 GGGGAAAGTC CAGGGCTATA GGGGTGTGATG AAGTCACCCC TTTCTGTOCC ACTACATGTG
1501 GGACTGACTT TCCGAGCCTC CAGTGTGTOCC CCATGGGTOCC CCTCCCCCTCT CAGCATTTCCT TGTCCCGTC
1561 CTGCATCTCA TGAGTGTGCC CCAATGCCAAGCT TTGGCTTATG GTTTTCATTC ATTGGTGAAG
1621 TGGACCTGGG GAGTGGTTAG GCAGCAAGCCT TTGAAGCCTTGT AAAGCCTGTG TTGAACAAGA TGTGGGGCCTT
1681 TAAATTAGGC TAGTTGCACGT AAAGCCTGTG GGTTTGGTCC TTGAACAAGA TGTGGGGCCTT
1741 GCAAGATGGGG AGAGTAAACCTTAAG
```

Figure 6: Ribonuclease Protection Assay

ds RNA-RNA (hybridized)

mRNA target A

RNA probe for mRNA A

ss RNA (not hybridized)

Nuclease treatment

Protected RNA species of predicted size, characteristic of mRNA A
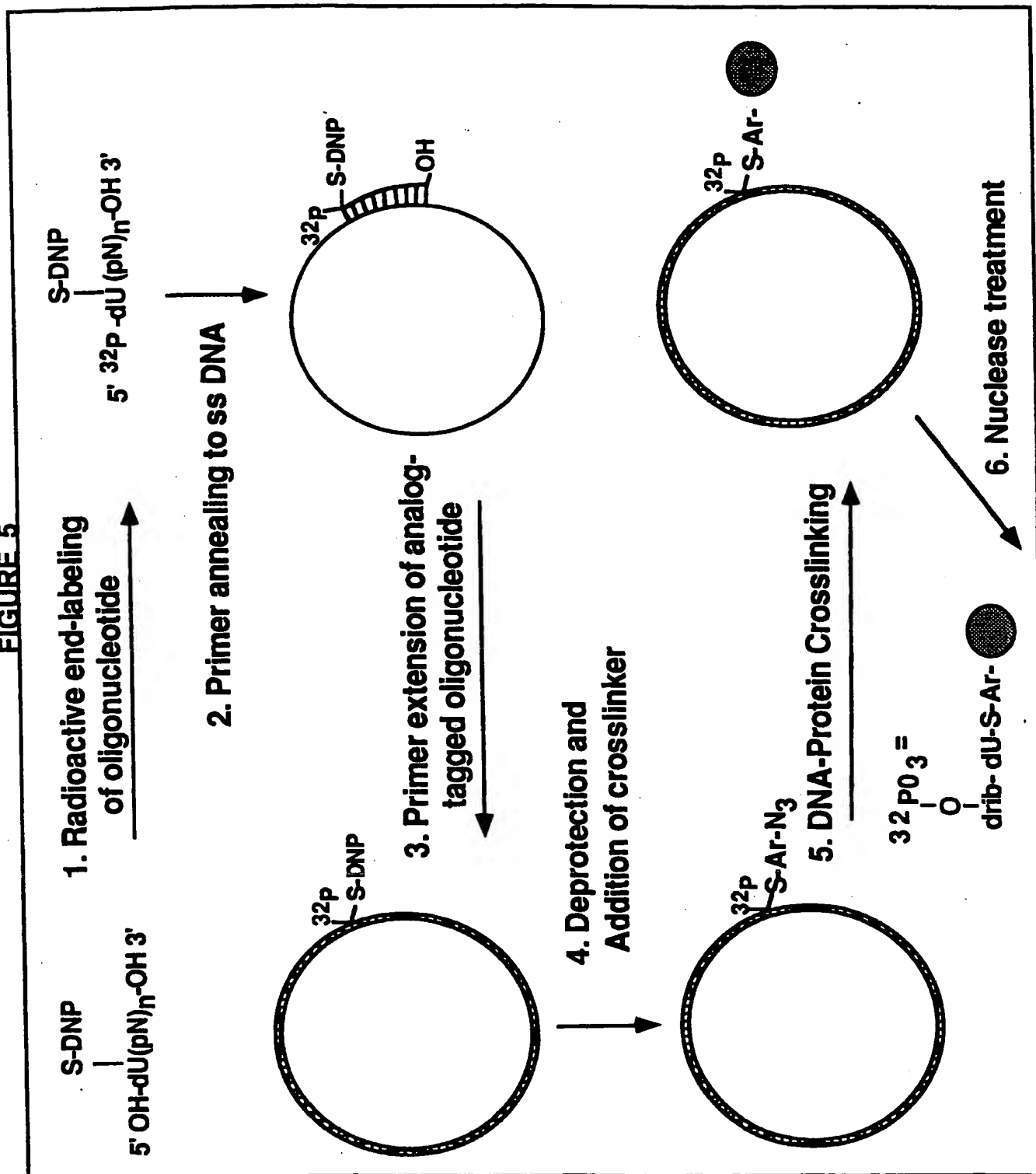
Hatched areas indicate regions of RNA-RNA homology with the target gene (hybridization and double-stranded structure). Shaded areas in the RNA probe are not homologous to the target mRNA and therefore do not hybridize, leaving regions of single-stranded RNA. These ssRNA regions are susceptible to cleavage by certain nucleases.
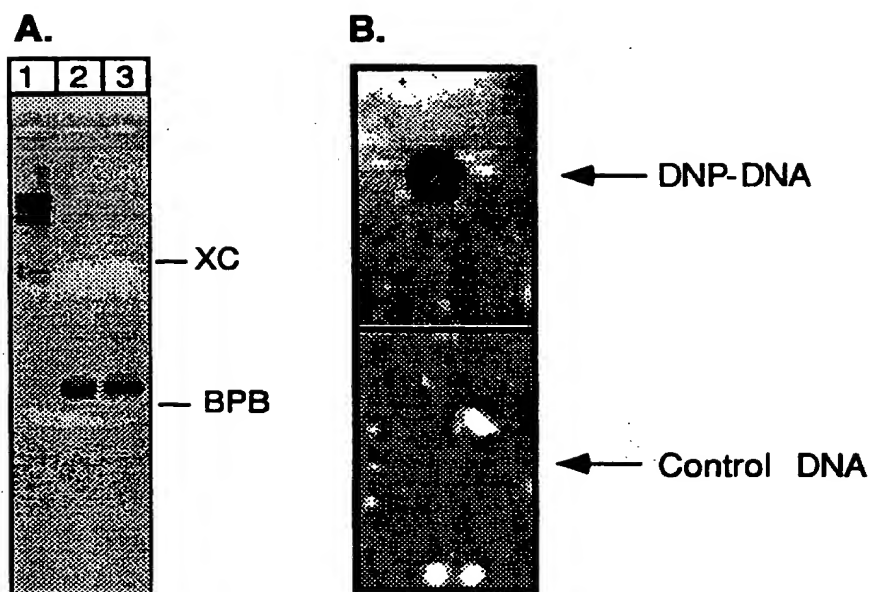
FIGURE 5

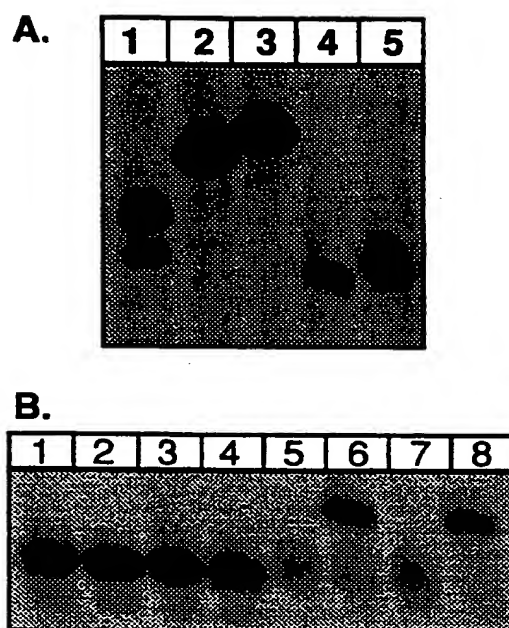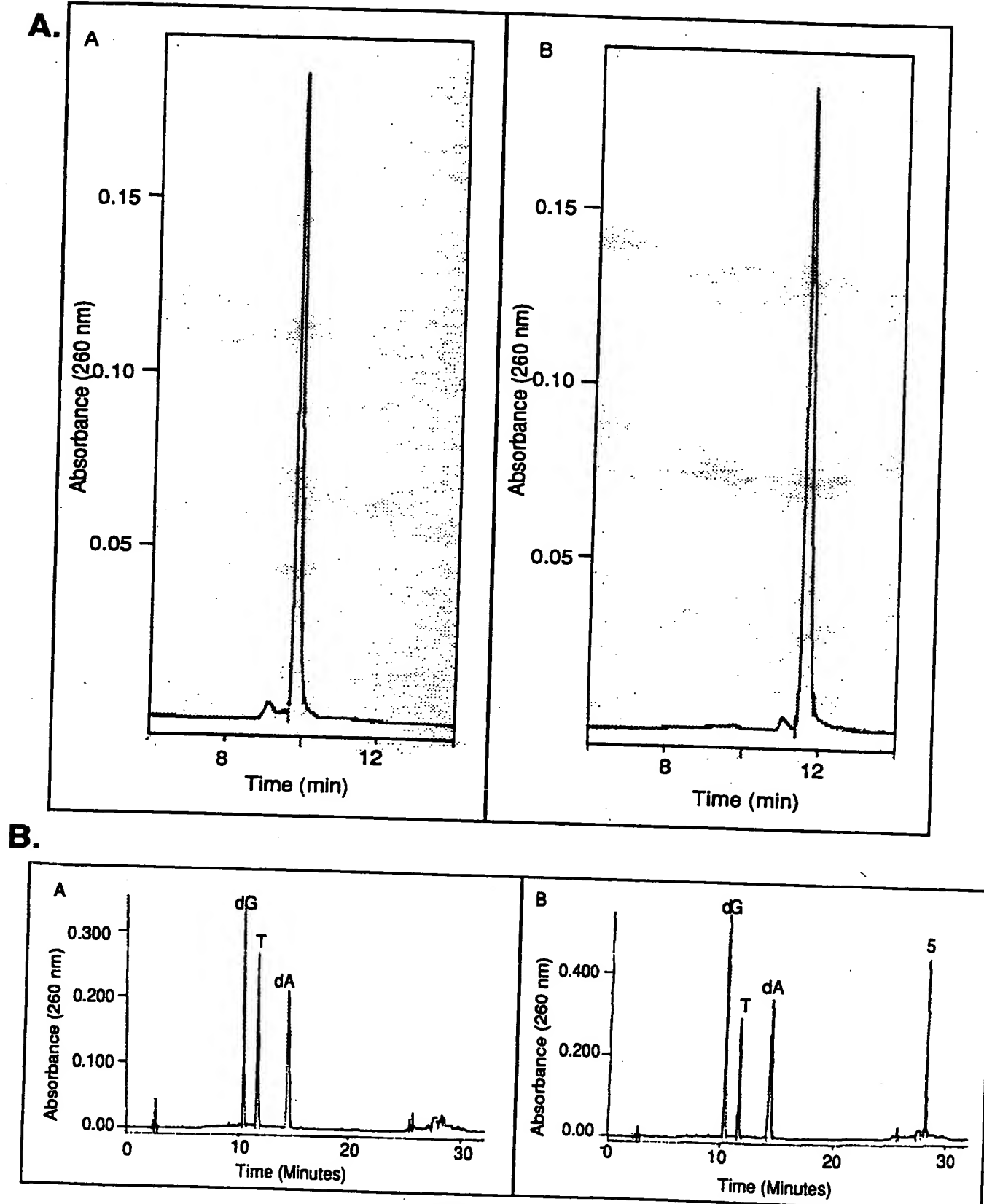1. Radioactive end-labeling of oligonucleotide

2. Primer annealing to ss DNA

3. Primer extension of analog-tagged oligonucleotide

4. Deprotection and Addition of crosslinker

5. DNA-Protein Crosslinking

6. Nuclease treatment

FIGURE 4

FIGURE 3

A.

B.

FIGURE 2

FIGURE 1